University of Alberta

WIRELESS CHANNEL STATE AND MODEL PARAMETER ESTIMATION

by

Yunfei Chen ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Department of Electrical and Computer Engineering

Edmonton, Alberta
Spring 2006

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-494-13950-1
Our file Notre référence
ISBN: 0-494-13950-1

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

For my lovely wife

Abstract

Practical wireless communication channels are usually characterized by models. The models are determined by parameters. They are also represented by realizations in specific observation intervals that determine channel state. This thesis contributes to the estimation of both channel state and model parameters.

First, maximum likelihood decision-based estimators for the channel state parameters are developed. Effects of channel estimation errors on the performances of two selection diversity combiners are evaluated. Novel diversity receivers using statistics of the channel estimation errors are designed. Optimum pilot symbol assisted modulation using pilot symbols for channel state parameter estimation is also investigated. As well, novel non-data-aided maximum likelihood estimators for the channel state parameters in an ultra-wide bandwidth system are derived, and the Cramér-Rao lower bounds are calculated analytically.

Second, maximum likelihood and moment-based estimators for the channel model parameters are proposed by using noisy channel samples. The estimators operate with or without knowledge of the noise power. Also, maximum likelihood estimators for the Ricean K parameter are derived by using fading phase samples, a method not considered previously.

Finally, maximum likelihood estimation of signal-to-noise ratio is studied. Two measures of signal-to-noise ratio are considered. The performances of the estimators are analyzed under the assumption of no decision errors. Using both known and unknown symbols in a frame, an approximate maximum likelihood estimator for signal-to-noise ratio is derived. A non-data-aided moment-based estimator for signal-to-interference-plus-noise ratio in a quadrature amplitude modulation system is also developed. In other works, maximum likelihood estimation of the average signal-to-noise ratio and a joint estimation of the K parameter and the average signal-to-noise ratio in a Ricean fading channel are performed. As the last part of this thesis, some concluding remarks are made and future possible works are outlined.

Acknowledgement

First, I would like to thank my thesis advisor Dr. Norman C. Beaulieu for his encouragement, enlightment, support and supervision. I have learned a lot from him. I am sure that this will benefit me throughout my career and my lifetime.

I would like to thank Dr. Chintha Tellambura, Dr. Julian Cheng and many other colleagues in the iCORE Wireless Communications Laboratory. Discussion with them is always a great pleasure to me. I am very fortunate to have them in the same laboratory. I would like to thank Dr. Ali Abdi, Dr. Peter Hooper, Dr. Masoud Ardakani and Dr. Nelson Durdle for agreeing to be my thesis examiners. Their kindness is greatly appreciated. I would also like to thank Dr. Rohana Karunamuni. His course in statistical inferences is one of the best and most useful courses I have ever taken.

I would like to thank my wife for her unselfish and continuing support. Life in a foreign country is not as easy as expected. She is always the sunshine that brightens my life. I am very lucky to have her with me. She is also one of the smartest persons I have ever met. She has a very solid understanding in mathematics. Her expertise helps me a lot.

I would like to thank my parents for their support. They persuaded me to pursue a doctoral degree in Canada. Their encouragement and vision always inspire me to do better.

Finally, I would like to thank the Alberta Ingenuity Fund and the Alberta Informatics Circle of Research Excellence (*iCORE*). Their financial support for my doctoral study is highly appreciated.

Table of Contents

1 Introduction **1**

1.1 Wireless Communication Channel 1

1.2 Importance of Wireless Channel Estimation 3

1.3 Methods of Wireless Channel Estimation 4

1.4 Performance Measures of Wireless Channel Estimators 6

2 Wireless Channel State Parameter Estimation **9**

2.1 ML Decision-Based Estimation of Channel State Parameter 10

2.1.1 Structures of ML Channel State Parameter Estimators 10

2.1.2 Performances of ML Channel State Parameter Estimators 14

2.2 SDC MFSK with Channel State Parameter Estimation Errors 20

2.2.1 System Model 21

2.2.2 $\hat{r}$-SDC 23

2.2.3 $\hat{\rho}$-SDC 28

2.2.4 Numerical Examples 32

2.3 Receiver Designs Using Error Statistics of Channel State Parameter Estimation . . 37

2.3.1 Case 1 39

2.3.2 Case 2 41

2.3.3 Comparison of Case 1 and Case 2 43

2.4 Receiver Designs Using Pilot Symbols of Channel State Parameter Estimation . . . 46

    2.4.1 System Model 46
    2.4.2 Optimum PSAM Signal Detectors 50
        2.4.2.1 Rayleigh Fading Channel 50
        2.4.2.2 Ricean Fading Channel 53
    2.4.3 Conventional PSAM Signal Detectors 55
    2.4.4 Comparison of Optimum and Conventional PSAM Signal Detectors 57
  2.5 Channel State Parameter Estimation in a UWB System 64
    2.5.1 UWB System Model 64
    2.5.2 CRLB for NDA ML UWB Channel Estimation 65
    2.5.3 Novel NDA ML UWB Channel State Parameter Estimators 68

3 Wireless Channel Model Parameter Estimation **76**
  3.1 Channel Model Parameter Estimation with Known Noise Power 76
    3.1.1 System Model 77
    3.1.2 Estimators for the Rayleigh Distribution Parameter 77
    3.1.3 Estimators for the Ricean Distribution Parameters 78
    3.1.4 Estimators for the Nakagami-*m* Distribution Parameters 80
    3.1.5 Numerical Results 81
  3.2 Channel Model Parameter Estimation with Unknown Noise Power 88
    3.2.1 System Model 88
    3.2.2 Estimators for the Rayleigh Distribution Parameter 89
    3.2.3 Estimators for the Ricean Distribution Parameters 91
    3.2.4 Estimators for the Nakagami-*m* Distribution Parameters 94
    3.2.5 Numerical Results 96
  3.3 Channel Model Parameter Estimation Using Phase Samples 101
    3.3.1 System Model 102
    3.3.2 K Estimator Using Phase Samples Only 103

3.3.3 K Estimator Using Both Phase Samples and Envelope Samples 104
3.3.4 Numerical Results 105

4 SNR Estimation in Wireless Communication Channel 108

4.1 ML Decision-Based Estimation of SNR 109
4.1.1 Static AWGN Channel 110
4.1.2 Slowly Fading Channel 113
4.1.3 Numerical Examples 116
4.2 ML Estimation of SNR Using Both Pilot and Data Symbols 123
4.2.1 Estimator Structure 124
4.2.2 Performance Analysis 125
4.3 ML Estimation of ASNR in a Ricean Channel 129
4.3.1 Estimator Structure 129
4.3.2 Analytical Performance 131
4.4 Joint Estimation of ASNR and K Parameter in a Ricean Channel 132
4.4.1 System Model 134
4.4.2 DA Estimation 136
4.4.2.1 Time-varying LOS Component 137
4.4.2.2 Constant LOS Component 140
4.4.3 NDA Estimation 141
4.4.3.1 Time-varying LOS Component 142
4.4.3.2 Constant LOS Component 143
4.4.4 Numerical Results 145
4.5 Moment-based Estimation of SINR 161
4.5.1 Channel Condition 1 161
4.5.2 Channel Condition 2 164

5 Conclusions and Future Work **167**

5.1 Conclusions 167

5.2 Future Work 169

References **170**

Appendix A **182**

Appendix B **184**

Appendix C **186**

Appendix D **187**

Appendix E **189**

List of Tables

2.1 Performance gains of OPSAM over CPSAM for BPSK 62

2.2 Performance gains of OPSAM over CPSAM for 16-QAM 63

List of Figures

2.1 Comparison of the performances of $\hat{r}$-SDC (maximum estimated signal amplitude) and r-SDC (maximum signal amplitude) for different values of q when $N_2 = 5N_1$ and NC-BFSK is used. 34

2.2 Comparison of the performances of $\hat{\rho}$-SDC (maximum estimated SNR) and ρ-SDC (maximum SNR) for different values of p when $q = 10$, $N_2 = 5N_1$ and NC-BFSK is used. 34

2.3 Comparison of the performances of $\hat{\rho}$-SDC (maximum estimated SNR) for selected values of p and q when $N_2 = 5N_1$ and NC-BFSK is used. 35

2.4 Comparison of the performances of $\hat{\rho}$-SDC (maximum estimated SNR) with different combinations of values of p and q when $p+q = 30$, $N_2 = 5N_1$ and NC-BFSK is used. 36

2.5 Performances of $\hat{r}$-SDC (maximum estimated signal amplitude), $\hat{\rho}$-SDC (maximum estimated SNR), r-SDC (maximum signal amplitude), and ρ-SDC (maximum SNR) at $p = 2$ and $q = 2$ with $N_2 = N_1$ for NC-BFSK signaling. 37

2.6 Performances of $\hat{r}$-SDC (maximum estimated signal amplitude), $\hat{\rho}$-SDC (maximum estimated SNR), r-SDC (maximum signal amplitude), and ρ-SDC (maximum SNR) at $p = 2$ and $q = 2$ with $N_2 = 5N_1$ for NC-BFSK signaling. 38

2.7 Performances of perfect MRC in (2.109), MRC with estimation errors in (2.110) and the new diversity receiver in (2.120) for BPSK signaling at $\lambda = 0.1$ (solid line) and $\lambda = 0.4$ (dash-dotted line) when $L_d = 2$ in Ricean fading channels. 44

2.8 Performances of perfect MRC in (2.109), MRC with estimation errors in (2.110) and the new diversity receiver in (2.120) for BPSK signaling at $\lambda = 0.1$ (solid line) and $\lambda = 0.4$ (dash-dotted line) when $L_d = 4$ in Ricean fading channels. 44

2.9 Performances of perfect MRC in (2.109), MRC with estimation errors in (2.110), the new diversity receivers in (2.115), (2.116) and (2.120) for 16-QAM signaling at $\lambda = 0.1$ (solid line) and $\lambda = 0.4$ (dash-dotted line) when $L_d = 2$ in Ricean fading channels. 45

2.10 Performances of perfect MRC in (2.109), MRC with estimation errors in (2.110), the new diversity receivers in (2.115), (2.116) and (2.120) for 16-QAM signaling at $\lambda = 0.1$ (solid line) and $\lambda = 0.4$ (dash-dotted line) when $L_d = 4$ in Ricean fading channels. 45

2.11 Performance comparison of the OPSAM signal detector and the CPSAM signal detector for BPSK signaling in Rayleigh and Ricean fading channels when $f_D T = 0.03$. 59

2.12 Performance comparison of the OPSAM signal detector and the CPSAM signal detector for BPSK signaling in Rayleigh and Ricean fading channels when $f_D T = 0.06$. 59

2.13 Performance comparison of the OPSAM signal detector and the CPSAM signal detector for BPSK signaling in Rayleigh and Ricean fading channels when $f_D T = 0.09$. 60

2.14 Performance comparison of the OPSAM signal detector and the CPSAM signal detector for 16-QAM signaling in Rayleigh and Ricean fading channels when $f_D T = 0.03$. 61

2.15 Performance comparison of the OPSAM signal detector and the CPSAM signal detector for 16-QAM signaling in Rayleigh and Ricean fading channels when $f_D T = 0.06$. 61

2.16 Performance comparison of the OPSAM signal detector and the CPSAM signal detector for 16-QAM signaling in Rayleigh and Ricean fading channels when $f_D T = 0.09$. 62

2.17 The square root of the CRLB for ML estimation of γ_1 at different values of ρ, when $\gamma_1 = 0.73$, $\gamma_2 = 0.67$ and $\gamma_3 = 0.35$. The CRLB for DA ML estimation is also shown. 68

2.18 The square root of the CRLB for ML estimation of τ_1 at different values of ρ, when $\gamma_1 = 0.73$, $\gamma_2 = 0.67$ and $\gamma_3 = 0.35$. The CRLB for DA ML estimation is also shown. 69

2.19 The square root of the CRLB for ML estimation of γ_1 at different values of γ_1, when $\rho = 0\,dB$. The CRLB for DA ML estimation is also shown. 69

2.20 The square root of the CRLB for ML estimation of τ_1 at different values of γ_1, when $\rho = 0\,dB$. The CRLB for DA ML estimation is also shown. 70

2.21 The normalized biases of τ_{NDA1}, τ_{NDA2} and τ_{DA} for different numbers of users. . . . 72

2.22 The normalized root mean squared errors of τ_{NDA1}, τ_{NDA2} and τ_{DA} for different numbers of users. 73

2.23 The normalized biases of γ_{NDA1}, γ_{NDA2} and γ_{DA} for different numbers of users. . . . 73

2.24 The normalized root mean squared errors of γ_{NDA1}, γ_{NDA2} and γ_{DA} for different numbers of users. 74

2.25 The normalized biases of γ_{NDA1} and γ_{DA} at small values of SINR. 74

2.26 The normalized root mean squared errors of γ_{NDA1} and γ_{DA} at small values of SINR. 75

3.1 Comparison of the normalized sample means of $\hat{K}_{MB1}$ given in (3.11) and $\hat{K}_{MB2}$ given in (3.13) with the true value for $L = 500$, $\sigma^2 = 0.1$ and $L = 500$, $\sigma^2 = 0.5$. . 82

3.2 Comparison of the normalized sample means of $\hat{K}_{MB1}$ given in (3.11) and $\hat{K}_{MB2}$ given in (3.13) with the true value for $L = 1000$, $\sigma^2 = 0.1$ and $L = 1000$, $\sigma^2 = 0.5$. 82

3.3 Comparison of the RMSEs of $\hat{K}_{MB1}$ given in (3.11) and $\hat{K}_{MB2}$ given in (3.13) with the CRLBs for $L = 500$, $\sigma^2 = 0.1$, $L = 500$, $\sigma^2 = 0.5$, $L = 1000$, $\sigma^2 = 0.1$ and $L = 1000$, $\sigma^2 = 0.5$. 83

3.4 Comparison of the normalized sample mean of $\hat{\Omega}_{Rice}$ given in (3.8a) with the true value for $L = 100$, $\sigma^2 = 0.1$, $L = 100$, $\sigma^2 = 0.5$, $L = 1000$, $\sigma^2 = 0.1$ and $L = 1000$, $\sigma^2 = 0.5$. 84

3.5 Comparison of the RMSE of $\hat{\Omega}_{Rice}$ given in (3.8a) with the CRLBs for $L = 100$, $\sigma^2 = 0.1$, $L = 100$, $\sigma^2 = 0.5$, $L = 1000$, $\sigma^2 = 0.1$ and $L = 1000$, $\sigma^2 = 0.5$. 84

3.6 Comparison of the normalized sample mean of $\hat{\Omega}_{Ray}$ given in (3.4) with the true value for $L = 100$, $\sigma^2 = 0.1$, $L = 100$, $\sigma^2 = 0.5$, $L = 1000$, $\sigma^2 = 0.1$ and $L = 1000$, $\sigma^2 = 0.5$. 85

3.7 Comparison of the RMSE of $\hat{\Omega}_{Ray}$ given in (3.4) with the CRLBs for $L = 100$, $\sigma^2 = 0.1$, $L = 100$, $\sigma^2 = 0.5$, $L = 1000$, $\sigma^2 = 0.1$ and $L = 1000$, $\sigma^2 = 0.5$. 85

3.8 Comparison of the normalized sample mean of $\hat{m}$ given in (3.16) with the true value for $L = 500$, $\sigma^2 = 0.1$, $L = 500$, $\sigma^2 = 0.5$, $L = 1000$, $\sigma^2 = 0.1$ and $L = 1000$, $\sigma^2 = 0.5$. 86

3.9 Comparison of the RMSE of $\hat{m}$ given in (3.16) with the CRLBs for $L = 500$, $\sigma^2 = 0.1$, $L = 500$, $\sigma^2 = 0.5$, $L = 1000$, $\sigma^2 = 0.1$ and $L = 1000$, $\sigma^2 = 0.5$. 86

3.10 Comparison of the normalized sample mean of $\hat{\Omega}_{Naka}$ given in (3.17) with the true value for $L = 100$, $\sigma^2 = 0.1$, $L = 100$, $\sigma^2 = 0.5$, $L = 1000$, $\sigma^2 = 0.1$ and $L = 1000$, $\sigma^2 = 0.5$. 87

3.11 Comparison of the RMSE of $\hat{\Omega}_{Naka}$ given in (3.17) with the CRLBs for $L = 100$, $\sigma^2 = 0.1$, $L = 100$, $\sigma^2 = 0.5$, $L = 1000$, $\sigma^2 = 0.1$ and $L = 1000$, $\sigma^2 = 0.5$. 88

3.12 Comparison of the normalized sample means of the estimators $\hat{K}_{MB1}$ in (3.38), $\hat{K}_{MB2}$ in (3.39), $\hat{K}'_{MB1}$ in (3.10) and $\hat{K}'_{MB2}$ in (3.12) with the true value for $L = 500$, TSNR = 7 dB (dashed line), $L = 500$, TSNR = 0 dB (dotted line) in a noisy Ricean fading channel. 96

3.13 Comparison of the normalized sample means of the estimators $\hat{K}_{MB1}$ in (3.38), $\hat{K}_{MB2}$ in (3.39), $\hat{K}'_{MB1}$ in (3.10) and $\hat{K}'_{MB2}$ in (3.12) with the true value for $L = 1000$, TSNR = 7 dB (solid line), $L = 1000$, TSNR = 0 dB (dashdotted line) in a noisy Ricean fading channel. 97

3.14 Comparison of the RMSEs of the estimators $\hat{K}_{MB1}$ in (3.38), $\hat{K}_{MB2}$ in (3.39), $\hat{K'}_{MB1}$ in (3.10) and $\hat{K'}_{MB2}$ in (3.12) with the CRLBs for $L = 500$, TSNR = 7 dB (dashed line), $L = 500$, TSNR = 0 dB (dotted line) in a noisy Ricean fading channel. 98

3.15 Comparison of the RMSEs of the estimators $\hat{K}_{MB1}$ in (3.38), $\hat{K}_{MB2}$ in (3.39), $\hat{K'}_{MB1}$ in (3.10) and $\hat{K'}_{MB2}$ in (3.12) with the CRLBs for $L = 1000$, TSNR = 7 dB (solid line), $L = 1000$, TSNR = 0 dB (dashdotted line) in a noisy Ricean fading channel. . 99

3.16 Comparison of the normalized sample means of $\hat{m}_{MB1}$ in (3.45), $\hat{m}_{MB2}$ in (3.46), $\hat{m}_s$ in (3.47) and $\hat{m}_t$ in (3.48) for $L = 500$, TSNR = 7 dB (dashed line), $L = 500$, TSNR = 0 dB (dotted line) in a noisy Nakagami-m fading channel. 100

3.17 Comparison of the normalized sample means of $\hat{m}_{MB1}$ in (3.45), $\hat{m}_{MB2}$ in (3.46), $\hat{m}_s$ in (3.47) and $\hat{m}_t$ in (3.48) for $L = 1000$, TSNR = 7 dB (solid line), $L = 1000$, TSNR = 0 dB (dashdotted line) in a noisy Nakagami-m fading channel. $\hat{m}_{old1}$ and $\hat{m}_{old2}$ use $L = 10,000$ and TSNR = 7 dB. 101

3.18 Comparison of the RMSEs of $\hat{m}_{MB1}$ in (3.45), $\hat{m}_{MB2}$ in (3.46), $\hat{m}_s$ in (3.47) and $\hat{m}_t$ in (3.48) with the CRLBs for $L = 500$, TSNR = 7 dB (dashed line), $L = 500$, TSNR = 0 dB (dotted line) in a noisy Nakagami-m fading channel. 102

3.19 Comparison of the RMSEs of $\hat{m}_{MB1}$ in (3.45), $\hat{m}_{MB2}$ in (3.46), $\hat{m}_s$ in (3.47) and $\hat{m}_t$ in (3.48) with the CRLBs for $L = 1000$, TSNR = 7 dB (solid line), $L = 1000$, TSNR = 0 dB (dashdotted line) in a noisy Nakagami-m fading channel. $\hat{m}_{old1}$ and $\hat{m}_{old2}$ use $L = 10,000$ and TSNR = 7 dB. 103

3.20 The biases of $\hat{K}_{ML}$, $\hat{K}_{AML}$ and $\hat{K}_{Env}$ for a sample size of $L = 500$ in a Ricean fading channel. 106

3.21 The RMSEs of $\hat{K}_{ML}$, $\hat{K}_{AML}$ and $\hat{K}_{Env}$ for a sample size of $L = 500$ in a Ricean fading channel. 106

3.22 The RMSEs of $\hat{K}_{ML}$, $\hat{K}_{AML}$ and $\hat{K}_{Env}$ versus the sample size with $K = 5$ in a Ricean fading channel. 107

4.1 Structures of the ML SNR estimators for ρ_1 and ρ_2 in an AWGN channel. 111

4.2 Structures of the ML SNR estimators for ρ_1 and ρ_2 in a slowly fading channel. . . 114

4.3 The biases of the sampled signal ML estimator for ρ_1 in an AWGN channel and ρ_1^R in a slowly fading channel at $I = 8$ (*) and $I = 20$ (o). 118

4.4 The RMSEs of the sampled signal ML estimator for ρ_1 in an AWGN channel and ρ_1^R in a slowly fading channel at $I = 8$ (*) and $I = 20$ (o). 118

4.5 The biases of the continuous time signal ML estimator for ρ_1 in an AWGN channel and ρ_1^R in a slowly fading channel at $BT = 8$ (*) and $BT = 20$ (o). 119

4.6 The RMSEs of the continuous time signal ML estimator for ρ_1 in an AWGN channel and ρ_1^R in a slowly fading channel at $BT = 8$ (*) and $BT = 20$ (o). 119

4.7 The biases of the sampled signal ML estimator for ρ_2 in an AWGN channel and a slowly fading channel at $I = 8$ (*) and $I = 20$ (o). 120

4.8 The RMSEs of the sampled signal ML estimator for ρ_2 in an AWGN channel and a slowly fading channel at $I = 8$ (*) and $I = 20$ (o). 120

4.9 The biases of the continuous time signal ML estimator for ρ_2 in an AWGN channel and a slowly fading channel at $BT = 8$ (*) and $BT = 20$ (o). 121

4.10 The RMSEs of the continuous time signal ML estimator for ρ_2 in an AWGN channel and a slowly fading channel at $BT = 8$ (*) and $BT = 20$ (o). 121

4.11 Comparison of the limiting Gaussian PDF and the PDF of the sampled signal ML estimator $\hat{\rho}_1$ in an AWGN channel for different values of sample sizes. 122

4.12 Comparison of the limiting Gaussian PDF and the PDF of the sampled signal ML estimator $\hat{\rho}_2$ in an AWGN channel for different values of sample sizes. 123

4.13 The biases of $\hat{\rho}$ (using both pilot symbols and data symbols), $\hat{\rho}_1$ (using pilot symbols only), and $\hat{\rho}_2$ (using data symbols only), when $L = 8$ and $Q = 28$. 126

4.14 The normalized root mean squared errors of $\hat{\rho}$ (using both pilot symbols and data symbols), $\hat{\rho}_1$ (using pilot symbols only), and $\hat{\rho}_2$ (using data symbols only), when $L = 8$ and $Q = 28$, and the Cramér-Rao lower bound. 126

4.15 The biases of the ML-based estimator $\hat{\rho}$, the M_2M_4 estimator and the SNV estimator when $L = 8$ and $Q = 28$. 128
4.16 The normalized root mean squared errors of the ML-based estimator $\hat{\rho}$, the M_2M_4 estimator and the SNV estimator when $L = 8$ and $Q = 28$, and the Cramér-Rao lower bound. 128
4.17 The biases (in dB) of $\hat{\hat{\rho}}$ for $I = 2$ and $K = 5$ for different sample sizes. 133
4.18 The normalized variances (with respect to the corresponding CRLBs) of $\hat{\hat{\rho}}$ for $I = 2$ and $K = 5$ for different sample sizes. 133
4.19 The normalized mean square errors (with respect to the corresponding CRLBs) of $\hat{\hat{\rho}}$ for $I = 2$ and $K = 5$ for different sample sizes. 134
4.20 The performance of $\hat{K}_{DA-LOSA}$ for different values of the sample lag with $L = 1024$ and $\bar{\rho} = 10\,dB$. 145
4.21 The performance of $\hat{K}_{DA-LOSB}$ for different values of the sample lag with $L = 1024$ and $\bar{\rho} = 10\,dB$. 146
4.22 The performance of $\hat{K}_{NDA-LOSA}$ for different values of the sample lag with $L = 1024$ and $\bar{\rho} = 10\,dB$. 146
4.23 The performance of $\hat{K}_{NDA-LOSB}$ for different values of the sample lag with $L = 1024$ and $\bar{\rho} = 10\,dB$. 147
4.24 The performance of $\hat{\hat{\rho}}_{DA-LOSA}$ for different values of the sample lag with $L = 1024$ and $K = 5$. 147
4.25 The performance of $\hat{\hat{\rho}}_{DA-LOSB}$ for different values of the sample lag with $L = 1024$ and $K = 5$. 148
4.26 The performance of $\hat{\hat{\rho}}_{NDA-LOSA}$ for different values of the sample lag with $L = 1024$ and $K = 5$. 148
4.27 The performance of $\hat{\hat{\rho}}_{NDA-LOSB}$ for different values of the sample lag with $L = 1024$ and $K = 5$. 149

4.28 The performance of $\hat{K}_{DA-LOSA}$ for mobile speed estimates having different accuracies at $L = 1024$, $\bar{\rho} = 10\,dB$. 150

4.29 The performance of $\hat{K}_{DA-LOSB}$ for mobile speed estimates having different accuracies at $L = 1024$ and $\bar{\rho} = 10\,dB$. 150

4.30 The performance of $\hat{K}_{NDA-LOSA}$ for mobile speed estimates having different accuracies at $L = 1024$ and $\bar{\rho} = 10\,dB$. 151

4.31 The performance of $\hat{K}_{NDA-LOSB}$ for mobile speed estimates having different accuracies at $L = 1024$ and $\bar{\rho} = 10\,dB$. 151

4.32 The performance of $\hat{\bar{\rho}}_{DA-LOSA}$ for mobile speed estimates having different accuracies at $L = 1024$ and $K = 5$. 152

4.33 The performance of $\hat{\bar{\rho}}_{DA-LOSB}$ for mobile speed estimates having different accuracies at $L = 1024$ and $K = 5$. 152

4.34 The performance of $\hat{\bar{\rho}}_{NDA-LOSA}$ for mobile speed estimates having different accuracies at $L = 1024$ and $K = 5$. 153

4.35 The performance of $\hat{\bar{\rho}}_{NDA-LOSB}$ for mobile speed estimates having different accuracies at $L = 1024$ and $K = 5$. 153

4.36 Performances of $\hat{K}_{DA-LOSB}$, $\hat{K}_{old}$ and $\hat{K}_{NDA-LOSB}$ with $L = 1024$ (dashed line) and $L = 512$ (solid line) at $\bar{\rho} = 10\,dB$, assuming a constant LOS fading component. . . 154

4.37 Performances of $\hat{K}_{DA-LOSB}$, $\hat{K}_{old}$ and $\hat{K}_{NDA-LOSB}$ with $L = 1024$ (dashed line) and $L = 512$ (solid line) at $\bar{\rho} = 5\,dB$, assuming a constant LOS fading component. . . . 155

4.38 Performances of $\hat{K}_{DA-LOSA}$, $\hat{K}_{old}$ and $\hat{K}_{NDA-LOSA}$ with $L = 1024$ (dashed line) and $L = 512$ (solid line) at $\bar{\rho} = 10\,dB$, assuming a time-varying LOS fading component. 156

4.39 Performances of $\hat{K}_{DA-LOSA}$, $\hat{K}_{old}$ and $\hat{K}_{NDA-LOSA}$ with $L = 1024$ (dashed line) and $L = 512$ (solid line) at $\bar{\rho} = 5\,dB$, assuming a time-varying LOS fading component. 156

4.40 Performances of $\hat{K}_{NDA-LOSA}$ at $\bar{\rho} = 5\,dB$ for different sample sizes, assuming a time-varying LOS fading component. 157

4.41 Performances of $\hat{\hat{\rho}}_{DA-LOSB}$, $\hat{\hat{\rho}}_{old}$ and $\hat{\hat{\rho}}_{NDA-LOSB}$ with $L = 1024$ (dashed line) and $L = 128$ (solid line) at $K = 5$, assuming a constant LOS fading component. 158

4.42 Performances of $\hat{\hat{\rho}}_{DA-LOSB}$, $\hat{\hat{\rho}}_{old}$ and $\hat{\hat{\rho}}_{NDA-LOSB}$ with $L = 1024$ (dashed line) and $L = 128$ (solid line) at $K = 2.5$, assuming a constant LOS fading component. . . . 158

4.43 Performances of $\hat{\hat{\rho}}_{DA-LOSA}$, $\hat{\hat{\rho}}_{old}$ and $\hat{\hat{\rho}}_{NDA-LOSA}$ with $L = 1024$ (dashed line) and $L = 128$ (solid line) at $K = 5$, assuming a time-varying LOS fading component. . . 160

4.44 Performances of $\hat{\hat{\rho}}_{DA-LOSA}$, $\hat{\hat{\rho}}_{old}$ in and $\hat{\hat{\rho}}_{NDA-LOSA}$ with $L = 1024$ (dashed line) and $L = 128$ (solid line) at $K = 2.5$, assuming a time-varying LOS fading component. . 160

4.45 The biases of $\hat{\rho}_1$ for 16-QAM and 256-QAM signals for different values of K. . . . 163

4.46 The root mean squared errors of $\hat{\rho}_1$ for 16-QAM and 256-QAM signals for different values of K. 164

4.47 The biases of $\hat{\rho}_2$ for 16-QAM and 256-QAM signals for different values of INR. . . 165

4.48 The root mean squared errors of $\hat{\rho}_2$ for 16-QAM and 256-QAM signals for different values of INR. 166

List of Acronyms

ALRT	Average likelihood ratio test
ASNR	Average signal-to-noise ratio
ASER	Average symbol error rate
AWGN	Additive white Gaussian noise
BER	Bit error rate
BPSK	Binary phase shift keying
CDF	Cumulative density function
CFSK	Coherent frequency shift keying
CPSAM	Conventional pilot symbol assisted modulation
CRLB	Cramér-Rao lower bound
DA	Data-aided
GLRT	Generalized likelihood ratio test
INR	Interference-to-noise ratio
LOS	Line-of-sight
MAI	Multiple access interference
MAP	Maximum *a posteriori*
MFSK	M-ary frequency shift keying
ML	Maximum likelihood
MM	Moment method
MMSE	Minimum mean squared error

MPSK	M-ary phase shift keying
MQAM	M-ary quadrature amplitude modulation
MRC	Maximal ratio combiner
MSE	Mean squared error
NC-BFSK	Non-coherent binary frequency shift keying
NCFSK	Non-coherent frequency shift keying
NDA	Non-data-aided
NRMSE	Normalized root mean squared error
OPSAM	Optimum pilot symbol assisted modulation
PDF	Probability density function
PSAM	Pilot symbol assisted modulation
QPSK	Quaternary phase-shift keying
RMSE	Root mean squared error
SDC	Selection diversity combiner
SER	Symbol error rate
SINR	Signal-to-interference-plus-noise ratio
SNR	Signal-to-noise ratio
SNV	Signal-to-noise variance
TSNR	Transmitted-signal-to-noise ratio
UWB	Ultra-wide bandwidth
16-QAM	16-ary quadrature amplitude modulation

List of Symbols

α^2	the variance of the fading gain
γ_l	the l-th multipath gain
$\delta(\cdot)$	the impulse function
θ	the fading phase
μ_n	the n-th order moment of a random variable
ρ	the signal-to-noise ratio
$\bar{\rho}$	the average signal-to-noise ratio
σ^2	the variance of the noise
τ_l	the l-th multipath delay
$\Gamma(\cdot)$	the Gamma function
Ω	the fading power
A	the channel gain
$\arctan(\cdot,\cdot)$	the arctangent function that returns a value over $(0,2\pi]$
$\cosh(\cdot)$	the hyperbolic cosine function
$Cov\{\cdot,\cdot\}$	the covariance
$D_v(\cdot)$	the hyperbolic cylinder function
$E\{\cdot\}$	the statistical mean
E	the signal energy
$erf(\cdot)$	the error function
$erfc(\cdot)$	the complementary error function

$F(\cdot,\cdot;\cdot;\cdot)$	the hypergeometric function
${}_1F_1(\cdot,\cdot;\cdot)$	the confluent hypergeometric function
$\tilde{f}_d$	the Doppler shift of the LOS component
f_D	the maximum Doppler shift
$I_n(\cdot)$	the n-th order modified Bessel function of the first kind
$Im\{\cdot\}$	the imaginary part of a complex number
$J_0(\cdot)$	the zero-th order Bessel function of the first kind
K	the Ricean K parameter
L	the sample size
L_c	the number of multipath components
L_d	the number of diversity branches
m	the Nakagami m parameter
M	the signaling constellation size
N_0	the single-sided power spectral density of the noise
N	the noise power
P^2	the local mean power of the LOS component
$Q_1(\cdot,\cdot)$	the Marcum's Q function
r	the fading envelope
$Re\{\cdot\}$	the real part of a complex number
$\operatorname{sech}(\cdot)$	the hyperbolic secant function
$\tanh(\cdot)$	the hyperbolic tangent function

Chapter 1

Introduction

Wireless communication has seen great development during recent years. It provides people with a way of talking to anyone, about anything, at any time in any place. To realize this goal, one of the most important techniques is wireless channel estimation. In wireless communication systems, signals sent over wireless channels are usually distorted by random channel variations. As a result, one has to obtain knowledge of the random channel distortions imposed on the signal in order to detect the transmitted signal as accurately as possible. In practice, this is often achieved by estimation. In this chapter, we first give a review of wireless communication channels. Then, we discuss the importance of wireless channel estimation by giving some examples. After this, we briefly introduce some frequently used methods of wireless channel estimation. Some performance measures for the wireless channel estimators are also presented.

1.1 Wireless Communication Channel

Generally speaking, there are two types of random distortions in wireless communication channels. One is fading, which is mainly caused by attenuations, scatterings and delays in the channel and is imposed on the signal multiplicatively. The other is noise, which is mainly caused by circuit electronics and is imposed on the signal additively. In a slowly and flatly fading channel, the fading is usually represented by a random variable $A = re^{j\theta}$ where r is the envelope and θ is the phase.

On the one hand, one may be interested in a gross description of the wireless channel distortions. This is often achieved by using channel models. For example, if there are scattering paths but no direct path between the transmitter and the receiver, the envelope of the fading, r, can be modeled as Rayleigh distributed. The probability density function (PDF) of the Rayleigh distribution is given by [1, eqn. (2.1-128)]

$$f_R(r) = \frac{2r}{\Omega} e^{-\frac{r^2}{\Omega}}, \quad r \geq 0 \tag{1.1}$$

where $\Omega = E\{r^2\}$ is the second moment and the only parameter that distinguishes different Rayleigh distributions. If there are scattering paths as well as a direct path between the transmitter and the receiver, r can be modeled as Ricean distributed. The PDF of a Ricean distribution is given by [1, eqn. (2.1-141)]

$$f_R(r) = \frac{r}{\alpha^2} e^{-\frac{r^2+P^2}{2\alpha^2}} I_0\left(\frac{rP}{\alpha^2}\right), \quad r \geq 0 \tag{1.2}$$

where P^2 is the local mean power of the direct path, $2\alpha^2$ is the local mean power of the random scattering paths and $I_0(\cdot)$ is the zero-th order modified Bessel function of the first kind [2, p. 374]. An alternative expression of (1.2) often used in the wireless literature is

$$f_R(r) = \frac{2(K+1)r}{\Omega} e^{-K-\frac{(K+1)r^2}{\Omega}} I_0\left(2\sqrt{\frac{K(K+1)}{\Omega}}r\right), \quad r \geq 0 \tag{1.3}$$

where $K = \frac{P^2}{2\alpha^2}$ is the K parameter measuring the relative strength of the direct path in the channel and $\Omega = E\{r^2\} = P^2 + 2\alpha^2$ is the second moment including both the local mean power of the direct path and the local mean power of the random scattering paths. The Ricean distribution is more flexible in modeling fading channels than the Rayleigh distribution in that it has two parameters. When $K = 0$ or $P^2 = 0$, there is no direct path and the fading envelope becomes Rayleigh distributed. A larger K corresponds to a better channel condition. In the extreme case when K approaches ∞, there is no fading in the channel. Another, similarly flexible fading envelope model, which also has two parameters is the Nakagami-m distribution. The PDF of the Nakagami-m distribution is given by [1, eqn. (2.1-147)]

$$f_R(r) = \frac{2r^{2m-1}}{\Gamma(m)} \left(\frac{m}{\Omega}\right)^m e^{-\frac{mr^2}{\Omega}}, \quad r \geq 0 \tag{1.4}$$

where $m = \frac{\Omega^2}{E\{(r^2-\Omega)^2\}}$ is the fading measure with $m \geq 0.5$ and $\Omega = E\{r^2\}$ is the second moment. The flexibility of this distribution derives from the fact that the m parameter can be chosen to model a wide range of fading conditions. For example, when $m = 0.5$, this is a one-sided Gaussian distribution representing very deep fading; when $m = 1$, the Nakagami-m distribution specializes to a Rayleigh distribution and when $m = \infty$, to a static (no fading) channel.

The phase of the fading and the additive noise can also be described by models. In most system analyses, the phase of the fading, θ, is modeled as uniformly distributed over $(0, 2\pi]$, which is the worst case that could ever happen. The noise is usually modeled as a zero-mean Gaussian random process, which is distinguished by its single-sided power spectral density denoted as N_0.

On the other hand, one can benefit more by having more detailed information about a particular wireless channel. This is often accomplished by measuring values of the random channel variations in specific observation intervals. These values, such as the values of r and θ, are crucial for some applications. As an example, the value of θ is required in coherent reception [1]. In this case, knowledge of the random channel distortions includes values determining the channel state as well.

1.2 Importance of Wireless Channel Estimation

In wireless communication systems, many techniques and components need channel state information, such as r and θ, or channel model information, such as N_0, m and K, to implement their algorithms. For example, knowledge of channel state information is required in order to calculate the branch metrics in the Viterbi decoder [1]. It is also needed in the likelihood ratio test in demodulators for higher-order modulations to achieve optimal detection [3]. Another application is diversity techniques. In diversity techniques, depending on what channel state knowledge is available, combiners of different performances can be implemented. Noteworthy, the optimal diversity combining performance is achieved only when the channel state is perfectly known [4].

In addition to channel state information, knowledge of channel model information is also indispensible in practical system designs. This knowledge is required in channel modeling where field measured data are used to determine the fading distribution. It is needed in some receiver designs, where the parameters of m and K are needed to make the data decisions [5], [6], [7]. In link budget calculation [8], adaptive modulation [9] and transmitter diversity optimization [10], these parameters are also required as they are good link quality measures.

Some applications may require both channel state and model information. One example is the measurement of signal-to-noise ratio (SNR). Since channel conditions in wireless communication systems are often changing with time, the SNR measure is a very important indicator of the channel quality. As a consequence, it is widely used in appliations, such as power control, error rate monitoring and transmission data rate adaptation, for the purpose of adaptive transmission [11], [12], [13]. The SNR is usually calculated as $\frac{r^2}{N_0}$. Thus, it is clear that both the channel state information r and the channel model information N_0 will be needed in order to derive the value of the SNR.

Due to the randomness of the channel variations, true values of the channel parameters are unknown in practical wireless communication systems. To provide these techniques and components with channel state and model information, wireless channel estimators have to be designed.

1.3 Methods of Wireless Channel Estimation

There are several methods of estimation, each having its own characteristics and applicabilities. The focus in this thesis will be placed on maximum likelihood (ML) estimation and moment method (MM) estimation.

The ML method is probably one of the most frequently used methods in communications. A ML estimator is derived as follows. Denote $f(x;a)$ as the PDF of the random variable X given the parameter a. The problem is to estimate the unknown parameter a using L observations of X, $X_1, X_2, \cdots, X_L$. Introduce the notation of the *likelihood function* as

$$f(a) = f(x_1, x_2, \cdots, x_L; a). \tag{1.5}$$

A ML estimate of a, $\hat{a}$, is defined as the value satisfying

$$f(\hat{a}) \geq f(a) \tag{1.6}$$

where a is any value of the parameter in its permissible range. In wireless communication, a number of PDFs of the random variable X occur in exponential form. A more convenient way of dealing with such PDFs is to use the natural logarithm of the likelihood function, that is, the *log-likelihood function*, since the natural logarithm is a strictly increasing function that will not affect the inequality in (1.6). In this case, $\hat{a}$ can also be defined as the value satisfying

$$\ln f(\hat{a}) \geq \ln f(a). \tag{1.7}$$

Assuming that the desired maximum of the likelihood function is within the permissible range of a and the first order derivative of $f(a)$ or $\ln f(a)$ is continuous, in practice, the ML estimate of a, $\hat{a}$, is often derived by differentiating $f(a)$ or $\ln f(a)$ with respect to a, setting the derivative equal to zero and solving the resulting equation for a. If there are more than one unknown parameters, the likelihood function and the log-likelihood function can be expressed as $f(\mathbf{a})$ and $\ln f(\mathbf{a})$, respectively, where $\mathbf{a} = [a_1 a_2 \cdots a_W]$ is a vector of W unknown parameters. The ML estimates of these parameters may then be found by differentiating $f(\mathbf{a})$ or $\ln f(\mathbf{a})$ with respect to each unknown parameter, setting the W partial derivatives equal to zero, and solving the resulting W equations jointly.

The ML method has been widely used owing to its elegant properties [14]. First, as long as an efficient estimator exists, the ML method will produce it [14]. Actually, as can be seen, the ML method has a systematic rule to generate an estimator, which is preferred in engineering practices. Second, the ML estimator has asymptotical optimality. It has been proved in [14] that a ML estimate converges to the true value of the parameter in probability when $L \to \infty$, that is, a ML estimate is a *consistent* estimate. Therefore, assuming that the limitation of $L \to \infty$ and the expectation of $\hat{a}$ are interchangable, the ML estimate is asymptotically unbiased when L is large. Also, it has been proved that the ML estimate is asymptotically efficient, which means that its variance asymptotically reaches the Cramér-Rao lower bound (CRLB) when L is large. Another useful property of the ML estimation is its invariance property [15]. The invariance property of ML estimation maintains that the ML estimate of a function of several unknown parameters is the value that is the same function of ML estimates of the unknown parameters. This property will be used later.

The ML estimator has many desirable properties. However, it is not always available. For example, if the resulting equation after differentiation is highly nonlinear, the solution and therefore, the ML estimator may not be derived. Even if the ML estimator is available, sometimes it may still be too complex to implement. In this case, one would rather use some simpler estimators. Moment-based estimators are some.

Unlike ML estimators, MM estimators don't have any asymptotical optimality in general. However, they are simple, and often yield good performances when the sample size L is large. Some ML estimators may require iterative algorithms to implement. In this case, MM estimators can also provide a good initial estimate. A MM estimator is derived as follows. Assume there are W unknown parameters $\mathbf{a} = [a_1 a_2 \cdots a_W]$. Also, denote the n-th order moment of the random variable X as μ_n. By calculating W different moments of X, one has

$$\mu_{(1)} = g_1(a_1, a_2, \cdots, a_W)$$

$$\mu_{(2)} = g_2(a_1, a_2, \cdots, a_W)$$

$$\cdots$$

$$\mu_{(W)} = g_W(a_1, a_2, \cdots, a_W)$$

where $\mu_{(n)}$ indexes the n-th of W different moments, not necessarily the n-th order moment of X, and $g_1, g_2, \cdots, g_W$ are W functions including $\mathbf{a}$ as parameters. Next, replacing μ_n with $\frac{1}{L}\sum_{l=1}^{L} x_l^n$ in the equations and solving them, moment-based estimators for $\mathbf{a}$ can then be obtained. An important part of the MM is to use the average of the n-th power of L observations to approximate the n-th order moment of X. Generally speaking, the lower the value of n is, the smaller the variance of $\frac{1}{L}\sum_{l=1}^{L} x_l^n$ will be.

1.4 Performance Measures of Wireless Channel Estimators

The goal of estimation is to obtain an estimate of the unknown parameter that approximates the true value as closely as possible. The closeness of the estimate is usually evaluated by several common measures [16]. Denote a and $\hat{a}$ as the true value and the estimate of the parameter, respectively. An estimate $\hat{a}$ is said to be *unbiased* if

$$E\{\hat{a}\} = a \tag{1.9}$$

where $E\{\hat{a}\}$ is the mean of the estimate. If (1.9) is not statisfied, the estimate is said to be *biased*. In this case, by shifting the estimate by its bias, an unbiased estimator may also be derived. The unbiasedness of the estimator will guarantee that the estimator gives the true value of the unknown parameter "on the average". But an unbiased estimator is not necessarily a good estimator. The goodness of the estimator is also evaluated by its deviation from the true value. This requires another important measure, the *mean squared error* (MSE). The MSE of an estimate is defined as

$$MSE\{\hat{a}\} = E\{(\hat{a}-a)^2\} \tag{1.10}$$

where a is the true value of the parameter. Note that, generally speaking, the variance of the estimator, defined as $Var\{\hat{a}\} = E\{(\hat{a}-E\{\hat{a}\})^2\}$, is different from the MSE. They satisfy the relation

$$MSE\{\hat{a}\} = Var\{\hat{a}\} + (a - E\{\hat{a}\})^2. \tag{1.11}$$

However, for an unbiased estimator, since $E\{\hat{a}\} = a$, the variance and the MSE will give the same value. Sometimes, one is also interested in the positive square roots of the variance and the MSE of the estimator. The positive square root of the variance is called *standard deviation*, while the positive square root of the MSE is called *root mean squared error* (RMSE).

To say that an estimator is good, one wants it not only to be unbiased, but also to have a variance as small as possible. The Cramér-Rao inequality states that, if $\hat{a}$ is any unbiased estimate of a, under certain assumptions, the variance of $\hat{a}$ satisfies [17]

$$Var\{\hat{a}\} \geq \left(E\left\{\left[\frac{\partial \ln f(x;a)}{\partial a}\right]^2\right\}\right)^{-1} \tag{1.12}$$

where $f(x;a)$ is the PDF of the random variable X given the parameter a. The value on the right hand side of the inequality is called the Cramér-Rao lower bound. The main task of finding a good wireless channel parameter estimator is to find an unbiased estimator that approaches its CRLB as closely as possible. Certainly, the optimum estimator (in the sense of minimum variance) is the one that reaches the CRLB. Its estimate is called an *efficient* estimate.

To obtain the estimator performance measures defined in (1.9) and (1.10) analytically, one needs to know the PDF of $\hat{a}$. However, this is not always possible as the estimate $\hat{a}$ can be a highly complicated function of the observations. Even when the PDF of $\hat{a}$ is available, it may still be difficult to calculate the integrations required by the expectation. In this case, a simulation method has to be used to examine the performance of the estimator, and the performance measures of the mean, the variance and the MSE will be replaced by the performance measures of the sample mean, the sample variance and the sample MSE, respectively. Denoting $\hat{a}_l$ as the estimate obtained in the l-th run of the simulation and L as the total number of runs, the sample mean, sample variance and sample MSE of the estimate can be calculated by using

$$\bar{\hat{a}} = \frac{1}{L}\sum_{l=1}^{L}\hat{a}_l, \tag{1.13a}$$

$$SVar\{\hat{a}\} = \frac{1}{L}\sum_{l=1}^{L}(\hat{a}_l - \bar{\hat{a}})^2, \tag{1.13b}$$

$$SMSE\{\hat{a}\} = \frac{1}{L}\sum_{l=1}^{L}(\hat{a}_l - a)^2, \tag{1.13c}$$

respectively. Note that, in some cases, the sample variance is calculated by dividing the sum with $L-1$, instead of L. They have little difference when L is large though. The sample standard deviation and the sample RMSE can be defined and calculated accordingly.

Chapter 2

Wireless Channel State Parameter Estimation

Channel state parameter estimation algorithms can be generally categorized into three classes: data-aided (DA), blind, and decision-based. The DA algorithm operates with knowledge of the transmitted data by sending a training sequence or pilot symbols. In this way, many classical estimation techniques, such as minimum mean squared error (MMSE), maximum *a posteriori* (MAP) and maximum likelihood (ML), can be used, and the resulting solutions are usually simple and optimal (given that the data are known) [17]. However, the DA method adds overhead to the system. It lowers the system throughput as well as costs extra power. In addition, the repeated transmission of this overhead will become untolerable when the channel changes too quickly. The blind algorithm assumes no knowledge of the transmitted data. It estimates the channel parameters solely from the received signal [1]. This saves overhead expense and maximizes system throughput. However, the blind estimator usually underperforms the DA estimator. It is usually more complex than the DA estimator as well. The decision-based algorithm is a compromise between the DA method and the blind method. The decision-based estimator uses data decisions as if they were known in the estimation. As a result, the decision-based estimator usually performs poorer than the DA estimator but better than the blind estimator, and it is usually more complicated than the DA estimator but simpler than the blind estimator. In this chapter, we first study the problem of ML decision-based estimation of the channel state parameter. Based on this study, we evaluate the effects of channel estimation errors on the performance of selection diversity combiners (SDCs) for M-ary frequency shift keying (MFSK) signals. Next, we design new diversity receivers using statistics of the channel estimation errors. We also design new receivers using pilot symbols for channel state parameter estimation directly, and therefore, we study the problem of optimum pilot symbol assisted modulation (PSAM). Finally, we discuss ultra-wide bandwidth (UWB) channel state parameter estimation.

2.1 ML Decision-Based Estimation of Channel State Parameter

Here, we focus on ML decision-based estimation of the channel state parameter. Both static additive white Gaussian noise (AWGN) and slowly fading memoryless channels are studied. The ML estimators are developed for sampled systems with bandlimited AWGN as well as continuous-time systems. The performances of the estimators are examined analytically by deriving approximations to the probability density functions (PDFs) of the ML estimates in error-free operation. For convenience, we consider ML decision-based estimation of the noise power as well, although the titles of this chapter and this section concern channel state parameter only.

2.1.1 Structures of ML Channel State Parameter Estimators

We consider M-ary memoryless digital signaling where the transmitter sends one of M signals, $s^{(j)}(t)$, corresponding to the j-th symbol, $j = 1,2,\cdots,M$. We assume that the signal is bandlimited to B Hz and is non-zero only over a time duration of T_d. Although it is not possible for a signal to be strictly bandlimited and strictly time-limited simultaneously, this is often achieved closely in practice by appropriately truncating the signal. For example, in bandwidth efficient signaling using raised-cosine pulse shaping, the signal is often truncated at $[-6T_s, 6T_s]$, where T_s is the symbol period [18, p. 289]. In this case, one has $T_d = 12T_s$. Assume that a sequence of L independent symbols are used for estimation and that the channel remains approximately constant during the estimation [19]- [27]. Also, assume that perfect synchronization is achieved such that no intersymbol interference occurs and the channel is memoryless [22]. The data sequence is sent over channels corrupted by AWGN. The noisy received data sequence is bandlimited by an ideal prefilter to remove out-of-band noise. This gives the filtered received signal as

$$y(t) = As^{(k)}(t) + n(t) \tag{2.1}$$

where $s^{(k)}(t)$ and $n(t)$ is the filtered transmitted data sequence and filtered noise, respectively, A is the unknown channel gain, and k is the k-th sequence of M^L possible transmitted squences [22]. The noise $n(t)$ in (2.1) is a complex bandlimited AWGN process with $E\{n(t)n^*(\tau)\} = N_0 \frac{\sin \pi B(t-\tau)}{\pi(t-\tau)}$ [1]. The value of N_0 is assumed unkown. The prefilter output, $y(t)$, is sampled at $f_s = B$. The sampled received data sequence is

$$y_i = As_i^{(k)} + n_i \tag{2.2}$$

where $i = 1,2,\cdots,I$ is the sample index and $s_i^{(k)}$ and n_i is the i-th sample of the transmitted data sequence and the noise, respectively. Assuming that the data sequence occupies a time duration of T, the total number of samples is $I = f_sT$ or $I = \frac{T}{\Delta t}$ where $\Delta t = \frac{1}{f_s}$ is the time spacing between samples. In the case when a symbol signal spans only one symbol period, $T = LT_d$. However, if bandwidth efficient signaling such as raised-cosine pulse shaping is used, $T < LT_d$. The noise samples, n_k, are Gaussian random variables each with mean zero and variance $N = BN_0$. Since $\sin(\pi B \frac{i}{f_s}) = 0$ for integer i, the noise samples are independent [28], [29]. The log-likelihood function can be written in terms of the sampled received data squence as [17]

$$\ln f(\mathbf{y}|A,N,k) = -I\ln(\pi N) - \frac{\sum_{i=1}^{I}[(y_i - As_i^{(k)})^*(y_i - As_i^{(k)})]}{N} \tag{2.3}$$

where $k = 1,2,\cdots,M^L$ and $\mathbf{y} = (y_1,y_2,\cdots,y_I)$. Eqn. (2.3) is the basis for ML estimation of the channel state parameter.

We first consider a static AWGN channel. This case assumes that the static channel phase distortion is perfectly compensated and was studied in [19]- [27]. Then, A is an unknown real constant. Denote $E_{sd}^{(k)} = \sum_{i=1}^{I}|s_i^{(k)}|^2$ for $k = 1,2,\cdots,M^L$. The log-likelihood function in (2.3) can be rewritten as

$$\ln f(\mathbf{y}|A,N,k) = -I\ln(\pi N) - \frac{1}{N}\sum_{i=1}^{I}|y_i|^2 - \frac{A^2}{N}E_{sd}^{(k)} + \frac{2A}{N}\sum_{i=1}^{I}Re\{y_i s_i^{(k)^*}\}. \tag{2.4}$$

By differentiating (2.4) with respect to A and N, setting the derivatives equal to zero, and solving the resulting equations, one has the ML estimator for A as

$$\hat{A} = \frac{\sum_{i=1}^{I} Re\{y_i s_i^{(\hat{k})*}\}}{E_{sd}^{(\hat{k})}} \tag{2.5}$$

and the ML estimator for N as

$$\hat{N} = \frac{1}{I}\sum_{i=1}^{I} |y_i|^2 - \frac{[\sum_{i=1}^{I} Re\{y_i s_i^{(\hat{k})*}\}]^2}{IE_{sd}^{(\hat{k})}} \tag{2.6}$$

where $\hat{k}$ is the sequence data decision that maximizes $h_1(\mathbf{y}|k) = \frac{1}{E_{sd}^{(k)}}\left[\sum_{i=1}^{I} Re\{y_i s_i^{(k)*}\}\right]^2$ for M^L sequences that have unequal energies and $h_2(\mathbf{y}|k) = \sum_{i=1}^{I} Re\{y_i s_i^{(k)*}\}$ for M^L sequences that have equal energies.

The results in (2.5) and (2.6) when specialized to M-ary phase shift keying (MPSK) agree with previous results obtained in [22]. They were derived here for a discrete time system with bandlimiting prefilter. In other cases, it is also of interest to study their continous time peers. First, this will allow us to examine the limiting case of the discrete time estimator and to see how the bandwidth affects the estimator performance. Second, this will give us approximate estimators for a system with a very wide bandwidth. Finally, this will also facilitate the evaluation of the performance of a continuous time system using channel estimates. The underlying unknown paramaters in a continuous time system are the channel gain A and the noise power N too. The corresponding results for a continuous time system without bandlimiting prefilter can be obtained by employing similar reasoning and techniques as those employed in the analysis of digital matched filters in [29]. Let the prefilter bandwidth B grow without bound. Then, the bandlimited system becomes non-bandlimited as the prefilter bandwidth increases. One has from (2.5) and (2.6) that

$$\hat{A} = \frac{\Delta t \sum_{i=1}^{I} Re\{y_i s_i^{(\hat{k})*}\}}{\Delta t \sum_{i=1}^{I} |s_i^{(\hat{k})}|^2} \tag{2.7}$$

$$\rightarrow \frac{\int_T Re\{y(t) s^{(\hat{k})*}(t)\}dt}{\int_T |s^{(\hat{k})}(t)|^2 dt} \tag{2.8}$$

and

$$\hat{N} = \frac{\Delta t}{T}\sum_{i=1}^{I} |y_i|^2 - \frac{[\Delta t \sum_{i=1}^{I} Re\{y_i s_i^{(\hat{k})*}\}]^2}{T\Delta t E_{sd}^{(\hat{k})}} \tag{2.9}$$

$$\rightarrow \quad \frac{1}{T}\left[\int_T |y(t)|^2 dt - \frac{\left[\int_T Re\{y(t)s^{(\hat{k})^*}(t)\}dt\right]^2}{\int_T |s^{(\hat{k})}(t)|^2 dt}\right] \tag{2.10}$$

where one notes that $\Delta t \rightarrow 0$ as $B \rightarrow \infty$ and the fundamental theorem of calculus [30] is applied to the Riemann sums in (2.7) and (2.9). In (2.8) and (2.10), the integrations are taken over the time duration of the entire data sequence, T. Denote $E_{sc}^{(k)} = \int_T |s^{(k)}(t)|^2 dt$ for $k = 1, 2, \cdots, M^L$ as the continuous time energies of the transmitted data sequence. The sequence data decision in (2.8) and (2.10), $\hat{k}$, is derived by selecting the value of k that maximizes $h_{1c}(y(t)|k) = \frac{1}{E_{sc}^{(k)}}\left[\int_T Re\{y(t)s^{(k)^*}(t)\}dt\right]^2$ for sequences of unequal energies and $h_{2c}(y(t)|k) = \int_T Re\{y(t)s^{(k)^*}(t)\}dt$ for sequences of equal energies. Next, we derive estimators for slowly fading channels.

Unlike in a static AWGN channel, we don't assume perfect compensation of the channel phase distortion in a slowly fading channel, as it is dynamically changing. Then, A is an unknown complex constant and the log-likelihood function in (2.3) can be rewritten as

$$\ln f(\mathbf{y}|A,N,k) = -I\ln(\pi N) - \frac{1}{N}\sum_{i=1}^{I}|y_i|^2 - \frac{|A|^2}{N}E_{sd}^{(k)} + \frac{2}{N}\sum_{i=1}^{I} Re\{A^* y_i s_i^{(k)^*}\}. \tag{2.11}$$

Again, by differentiating (2.11) with respect to A and N, setting the derivatives equal to zero and solving the resulting equations, one has the ML estimators for A and N as

$$\hat{A} = \frac{\sum_{i=1}^{I} y_i s_i^{(\hat{k})^*}}{E_{sd}^{(\hat{k})}} \tag{2.12}$$

and

$$\hat{N} = \frac{1}{I}\sum_{i=1}^{I}|y_i|^2 - \frac{|\sum_{i=1}^{I} y_i s_i^{(\hat{k})^*}|^2}{I E_{sd}^{(\hat{k})}} \tag{2.13}$$

where $\hat{k}$ is the sequence data decision maximizing $h_3(\mathbf{y}|k) = \frac{1}{E_{sd}^{(k)}}\left|\sum_{i=1}^{I} y_i s_i^{(k)^*}\right|^2$. Note that this sequence detector is not applicable for all signalings, as the maximum of $h_3(\mathbf{y}|k)$ may have multiple solutions for signals such as MPSK.

By applying similar reasonings and techniques to those used in the previous section, results for continuous time are derived as

$$\hat{A} \quad = \quad \frac{\Delta t \sum_{i=1}^{I} y_i s_i^{(\hat{k})^*}}{\Delta t E_{sd}^{(\hat{k})}}$$

$$\rightarrow \quad \frac{\int_T y(t) s^{(\hat{k})^*}(t) dt}{E_{sc}^{(\hat{k})}} \tag{2.14}$$

and

$$\begin{aligned}
\hat{N} &= \frac{\Delta t}{T} \sum_{i=1}^{I} |y_i|^2 - \frac{|\Delta t \sum_{i=1}^{I} y_i s_i^{(\hat{k})^*}|^2}{T \Delta t E_{sd}^{(\hat{k})}} \\
&\rightarrow \frac{1}{T} \left[\int_T |y(t)|^2 dt - \frac{\left| \int_T y(t) s^{(\hat{k})^*}(t) dt \right|^2}{\int_T |s^{(\hat{k})}(t)|^2 dt} \right]
\end{aligned} \tag{2.15}$$

where the sequence decision $\hat{k}$ may be derived by selecting the value of k maximizing $h_{3_c}(y(t)|k) = \frac{1}{E_{sc}^{(k)}} \left| \int_T y(t) s^{(k)^*}(t) dt \right|^2$.

2.1.2 Performances of ML Channel State Parameter Estimators

Assuming that there are no decision errors, we have $\hat{k} = k$, where k is the index of the transmitted data sequence. Again, we begin with a static AWGN channel.

By using $\hat{k} = k$ and (2.2) in (2.5) and (2.6), one has

$$\hat{A} = A + \frac{\sum_{i=1}^{I} Re\{n_i s_i^{(k)^*}\}}{E_{sd}^{(k)}} \tag{2.16}$$

and

$$\hat{N} = \frac{1}{I} \sum_{i=1}^{I} |n_i|^2 - \frac{1}{I} \left(\frac{1}{\sqrt{E_{sd}^{(k)}}} \sum_{i=1}^{I} Re\{n_i s_i^{(k)^*}\} \right)^2. \tag{2.17}$$

From (2.16), $\hat{A}$ is a Gaussian random variable and its PDF is

$$p_{\hat{A}}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-A)^2}{2\sigma^2}} \tag{2.18}$$

where $\sigma^2 = \frac{N}{2E_{sd}^{(k)}}$. It is clear from (2.18) that this estimator is unbiased in the absence of decision errors. The corresponding Cramér-Rao lower bound (CRLB) for estimation of A in this case is $\frac{N}{2E_{sd}^{(k)}}$. Therefore, the sampled signal ML estimator for A in a static AWGN channel achieves the CRLB, and it is optimal in the sense of minimum variance in error-free operation. The PDF of $\hat{N}$ in (2.17) can be derived as follows. Denote

$$X = \frac{1}{\sqrt{E_{sd}^{(k)}}} \sum_{i=1}^{I} Re\{n_i s_i^{(k)^*}\} \tag{2.19}$$

$$
\begin{aligned}
Y &= \sum_{i=1}^{I} |n_i|^2 && (2.20)\\
Z &= Y - X^2. && (2.21)
\end{aligned}
$$

It can be proved that X and Z are independent; a proof is given in Appendix A. Therefore, one has $\Psi_Y(jv) = \Psi_Z(jv)\Psi_{X^2}(jv)$, where $\Psi_Y(jv)$, $\Psi_Z(jv)$ and $\Psi_{X^2}(jv)$ are the characteristic functions of Y, Z and X^2, respectively. The characteristic functions of X^2 and Y are derived from (2.19) and (2.20) as [1, eqn. (2.1-109)]

$$\Psi_{X^2}(jv) = \frac{1}{(1 - jvN)^{\frac{1}{2}}} \tag{2.22}$$

and

$$\Psi_Y(jv) = \frac{1}{(1 - jvN)^{I}}. \tag{2.23}$$

Then, the characteristic function of Z is

$$\Psi_Z(jv) = \frac{1}{(1 - jvN)^{\frac{2I-1}{2}}} \tag{2.24}$$

where (2.22) and (2.23) are used. Finally, as $\hat{N} = \frac{Z}{I}$, the PDF of $\hat{N}$ is derived from (2.24) as

$$p_{\hat{N}}(x) = \frac{1}{\Gamma(I - \frac{1}{2})} \left(\frac{I}{N}\right)^{(I-\frac{1}{2})} x^{(I-\frac{3}{2})} e^{-\frac{Ix}{N}}, \quad x \geq 0 \tag{2.25}$$

which is a central chi-square distribution with $2I - 1$ degrees of freedom and where $\Gamma(\cdot)$ is the Gamma function [1, eqn. (2-1-110)]. The mean and the variance of $\hat{N}$ are $E\{\hat{N}\} = \frac{2I-1}{2I}N$ and $Var\{\hat{N}\} = \frac{2I-1}{2I^2}N^2$. The CRLB for estimation of N in this case is $\frac{N^2}{I}$. Therefore, the sampled signal ML estimator for N in a static AWGN channel is asymptotically optimal when I is large in error-free operation. An unbiased estimator when no decision errors occur is obtained by scaling $\hat{N}$ in (2.17) with $\frac{2I}{2I-1}$. This scaling will increase the variance of the estimate by a factor of $\left(\frac{2I}{2I-1}\right)^2$.

The PDFs of the continuous time ML estimates of A and N in a static AWGN channel can be derived similarly, assuming that no decision errors occur. Applying $\hat{k} = k$ and (2.1) in (2.8) and (2.10), one has

$$\hat{A} = A + C_1 \tag{2.26}$$

and

$$\hat{N} = C_2 - C_3 \tag{2.27}$$

with

$$C_1 = \frac{\int_T Re\{n(t)s^{(k)*}(t)\}dt}{E_{sc}^{(k)}} \tag{2.28a}$$

$$C_2 = \frac{\int_T |n(t)|^2 dt}{T} \tag{2.28b}$$

$$C_3 = \frac{\left[\int_T Re\{n(t)s^{(k)*}(t)\}dt\right]^2}{TE_{sc}^{(k)}}. \tag{2.28c}$$

The random variable C_1 in (2.28a) has a Gaussian distribution with mean zero and variance $\tilde{\sigma}^2 = \frac{N_0}{2E_{sc}^{(k)}}$, and the random variable C_3 in (2.28c) has a central chi-square distribution with one degree of freedom and parameter $\frac{N_0}{2T}$. Therefore, the ML estimate of A has a Gaussian random error C_1, while the ML estimate of N has a chi-square random error C_3 in the absence of decision errors, as C_2 in (2.28b) is the noise power in the T-second interval. Since C_1 is a Gaussian random variable, from (2.26), the continuous time ML estimate of A, $\hat{A}$, has a Gaussian PDF

$$p_{\hat{A}}(x) = \frac{1}{\sqrt{2\pi\tilde{\sigma}^2}} e^{-\frac{(x-A)^2}{2\tilde{\sigma}^2}} \tag{2.29}$$

with mean A and variance $\tilde{\sigma}^2$ when no decision errors occur. The PDF of $\hat{N}$ can be derived by using similar reasoning to that used to derive (2.25). Since C_3 is a central chi-square random variable, its characteristic function is

$$\Psi_{C_3}(jv) = \frac{1}{(1 - jv\frac{N_0}{T})^{\frac{1}{2}}}. \tag{2.30}$$

The quantity C_2 is the time-average power of $n(t)$ in a T-second interval. The distribution of integrated squared Gaussian noise is unknown; in fact, it is a long-standing problem to find its distribution in communication theory and applied physics [31], [32]. A useful closed-form approximation to the PDF was derived in [33]. Applying the trapezoidal rule of integration to the integral in C_2 given by (2.28b), one has [33]

$$C_2 \approx \tilde{C}_2 = R_1^2 + R_2^2 \tag{2.31}$$

where $R_1^2 = \frac{|n(0)|^2}{2J} + \frac{|n(T)|^2}{2J}$, $R_2^2 = \frac{1}{J}\sum_{j=1}^{J-1} |n(\frac{jT}{J})|^2$, and $J+1$ is the number of samples used in the numerical integration. Setting $J = BT$, one has $J+1$ independent complex Gaussian samples $\left\{n(\frac{jT}{J})\right\}_{j=0}^{J}$ [33]. Therefore, R_1^2 is a Gamma random variable with characteristic function $\Psi_{R_1^2}(jv) = \frac{1}{(1-jv\frac{N_0}{2T})^2}$ and R_2^2 is a Gamma random variable with characteristic function $\Psi_{R_2^2}(jv) = \frac{1}{(1-jv\frac{N_0}{T})^{J-1}}$. Since R_1^2 and R_2^2 are independent, one has [34, eqn. (6-193)]

$$\Psi_{\tilde{C}_2}(jv) = \frac{1}{(1-jv\frac{N_0}{2T})^2(1-jv\frac{N_0}{T})^{J-1}} \tag{2.32}$$

Finally, as $C_2 = \hat{N} + C_3$ from (2.27), $\tilde{C}_2 = \hat{N}_A + C_3$ where $\hat{N}_A \approx \hat{N}$ is an approximation to $\hat{N}$. Then, $\Psi_{\tilde{C}_2}(jv) = \Psi_{\hat{N}_A}(jv)\Psi_{C_3}(jv)$. Using (2.30) and (2.32), one has

$$\Psi_{\hat{N}_A}(jv) = \frac{1}{(1-jv\frac{N_0}{2T})^2(1-jv\frac{N_0}{T})^{J-\frac{3}{2}}}. \tag{2.33}$$

The PDF of the approximation to $\hat{N}$, $\hat{N}_A$, is derived from (2.33) by using [35, eqn. (3.384.7)] as

$$p_{\hat{N}_A}(x) = \frac{4}{\Gamma(BT+\frac{1}{2})}\left(\frac{BT}{N}\right)^{BT+\frac{1}{2}} x^{BT-\frac{1}{2}} e^{-\frac{BT}{N}x} {}_1F_1\left(2, BT+\frac{1}{2}; -\frac{BT}{N}x\right), \quad x \geq 0, BT > \frac{1}{2} \tag{2.34}$$

where ${}_1F_1(\cdot,\cdot;\cdot)$ is the confluent hypergeometric function [2, p. 504] and the relation $N = BN_0$ is used. The n-th order moment of $\hat{N}_A$ can be derived from (2.34) as

$$E\{\hat{N}_A^n\} = \frac{4}{2^{BT+n+\frac{1}{2}}} \cdot \frac{\Gamma(BT+n+\frac{1}{2})}{\Gamma(BT+\frac{1}{2})} \cdot \left(\frac{N}{BT}\right)^n \cdot F(BT-\frac{3}{2}, BT+n+\frac{1}{2}; BT+\frac{1}{2}; \frac{1}{2}) \tag{2.35}$$

with $BT > \frac{1}{2}$, $BT+n+\frac{1}{2} > 0$ and where $F(\cdot,\cdot;\cdot;\cdot)$ is the hypergeometric function [2, p. 556]. The mean and variance of $\hat{N}_A$ are obtained from (2.35) as $E\{\hat{N}_A\} = \frac{2BT-1}{2BT}N$ and $Var\{\hat{N}_A\} = \frac{BT-1}{(BT)^2}N^2$. An approximately unbiased estimator for N can be obtained by scaling $\hat{N}_A$ with $\frac{2BT}{2BT-1}$, which increases the variance as before.

Next, we analyze the performances of the estimators for a slowly fading channel. Denote $A = A_R + jA_I$, $n_R = \sum_{i=1}^{I} Re\{n_i s_i^{(k)*}\}$ and $n_I = \sum_{i=1}^{I} Im\{n_i s_i^{(k)*}\}$. By using $\hat{k} = k$ and (2.2) in (2.12) and (2.13), one has

$$\hat{A} = \hat{A}_R + j\hat{A}_I \tag{2.36}$$

and

$$\hat{N} = \frac{1}{I}\sum_{i=1}^{I} |n_i|^2 - \frac{1}{IE_{sd}^{(k)}} n_R^2 - \frac{1}{IE_{sd}^{(k)}} n_I^2 \tag{2.37}$$

where $\hat{A}_R = A_R + \frac{n_R}{E_{sd}^{(k)}}$ and $\hat{A}_I = A_I + \frac{n_I}{E_{sd}^{(k)}}$ are the ML estimates of the real and imaginary components of the complex channel gain, respectively. Inspection of the estimators $\hat{A}_R$ and $\hat{A}_I$ shows that each has the same performance as that of the estimator in (2.16). Their PDF's are then derived as

$$p_{\hat{A}_R}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-A_R)^2}{2\sigma^2}} \tag{2.38}$$

and

$$p_{\hat{A}_I}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-A_I)^2}{2\sigma^2}} \tag{2.39}$$

where $\sigma^2 = \frac{N}{2E_{sd}^{(k)}}$. Since the ML estimator for A in (2.36) is derived by combining $\hat{A_R}$ and $\hat{A_I}$, one has that $\hat{A}$ in (2.36) has mean A and variance $\frac{N}{E_{sd}^{(k)}}$. Since the CRLB for estimation of A in a slowly fading channel is $\frac{N}{E_{sd}^{(k)}}$, the sampled signal ML estimator for A in a slowly fading channel is optimal in the sense of minimum variance when no decision errors occur. The PDF of $\hat{N}$ in (2.37) can also be derived by using a characteristic function approach, as previously. Denote

$$S_1 = \frac{1}{\sqrt{E_{sd}^{(k)}}} n_R \tag{2.40}$$

$$S_2 = \frac{1}{\sqrt{E_{sd}^{(k)}}} n_I \tag{2.41}$$

$$U = Y - S_1^2 - S_2^2. \tag{2.42}$$

It is proved that the random variables S_1, S_2 and U are independent; a proof is given in Appendix B. Therefore, S_1^2, S_2^2 and U are also independent. Their characteristic functions satisfy $\Psi_Y(jv) = \Psi_U(jv)\Psi_{S_1^2}(jv)\Psi_{S_2^2}(jv)$, where $\Psi_Y(jv)$, $\Psi_{S_1^2}(jv)$, $\Psi_{S_2^2}(jv)$ and $\Psi_U(jv)$ are the characteristic functions of Y, S_1^2, S_2^2 and U, respectively. Note that $\Psi_Y(jv)$ can be derived from (2.23), while $\Psi_{S_1^2}(jv)$ and $\Psi_{S_2^2}(jv)$ can be derived from (2.22). One has

$$\Psi_U(jv) = \frac{1}{(1 - jvN)^{I-1}}. \tag{2.43}$$

Finally, as $\hat{N} = \frac{U}{I}$, one has from (2.43)

$$p_{\hat{N}}(x) = \frac{1}{\Gamma(I-1)} \left(\frac{I}{N}\right)^{(I-1)} x^{(I-2)} e^{-\frac{Ix}{N}}, \quad x \geq 0 \tag{2.44}$$

which is a central chi-square distribution with $2I-2$ degrees of freedom. Its mean and variance are $E\{\hat{N}\} = \frac{I-1}{I}N$ and $Var\{\hat{N}\} = \frac{I-1}{I^2}N^2$. The CRLB for estimation of N in a slowly fading channel is $\frac{N^2}{I}$. Thus, the sampled signal ML estimator for N in slowly fading channels is asymptotically optimal when I is large. An unbiased estimator can also be obtained by scaling the estimator in (2.37) with $\frac{I}{I-1}$, however, increasing the estimator variance.

By using $\hat{k} = k$ and (2.1) in (2.14) and (2.15), the continuous time ML estimators for A and N in a slowly fading channel can be rewritten in forms similar to those in (2.26) and (2.27), except that

$$C_1 = \frac{\int_T n(t) s^{(k)*}(t)dt}{E_{sc}^{(k)}} \tag{2.45a}$$

$$C_3 = \frac{\left|\int_T n(t) s^{(k)*}(t)dt\right|^2}{T E_{sc}^{(k)}} \tag{2.45b}$$

in this case. Similarly, the continuous time ML estimates of the real and imaginary components of A in a slowly fading channel, $\hat{A}_R$ and $\hat{A}_I$, have PDFs

$$p_{\hat{A}_R}(x) = \frac{1}{\sqrt{2\pi\tilde{\sigma}^2}} e^{-\frac{(x-A_R)^2}{2\tilde{\sigma}^2}} \tag{2.46}$$

and

$$p_{\hat{A}_I}(x) = \frac{1}{\sqrt{2\pi\tilde{\sigma}^2}} e^{-\frac{(x-A_I)^2}{2\tilde{\sigma}^2}} \tag{2.47}$$

where $\tilde{\sigma}^2$ is defined as before. Therefore, the ML estimate $\hat{A}$ is a complex Gaussian random variable with mean A and variance $2\tilde{\sigma}^2$ in error-free operation. By using similar techniques to those used to derive (2.34), the PDF of the approximation to $\hat{N}$, $\hat{N}_A$, can also be obtained. In a slowly fading channel, C_3 in (2.45b) has the characteristic function

$$\Psi_{C_3}(jv) = \frac{1}{(1 - jv\frac{N_0}{T})}. \tag{2.48}$$

Since $\Psi_{\tilde{C}_2}(jv) = \Psi_{\hat{N}_A}(jv)\Psi_{C_3}(jv)$ and $\Psi_{\tilde{C}_2}(jv)$ is given by (2.32), one has

$$\Psi_{\hat{N}_A}(jv) = \frac{1}{(1 - jv\frac{N_0}{2T})^2(1 - jv\frac{N_0}{T})^{J-2}} \tag{2.49}$$

and

$$p_{\hat{N}_A}(x) = \frac{4}{\Gamma(BT)} \left(\frac{BT}{N}\right)^{BT} x^{BT-1} e^{-\frac{BT}{N}x} {}_1F_1\left(2, BT; -\frac{BT}{N}x\right), \quad x \geq 0, BT > 1. \tag{2.50}$$

The n-th order moment of $\hat{N}_A$ in a slowly fading channel is

$$E\{\hat{N}_A^n\} = \frac{4}{2^{BT+n}} \cdot \frac{\Gamma(BT+n)}{\Gamma(BT)} \cdot \left(\frac{N}{BT}\right)^n \cdot F(BT-2, BT+n; BT; \frac{1}{2}) \tag{2.51}$$

with $BT > 1$ and $BT + n > 0$. The mean and the variance of $\hat{N}_A$ are $E\{\hat{N}_A\} = \frac{BT-1}{BT}N$ and $Var\{\hat{N}_A\} = \frac{BT-\frac{3}{2}}{(BT)^2}N^2$. By scaling $\hat{N}_A$ with $\frac{BT}{BT-1}$, one also has an approximately unbiased estimator with increased variance.

The above results are based on the assumption of no decision errors. They are valid for DA estimation and decision-based estimation with moderate to large values of signal-to-noise (SNR). For decision-based estimation at small values of SNR, they serve as benchmarks. Note that the system model used in the derivation is fairly simple, as our final goal is to estimate SNR in Chapter 4. Next, we evaluate the effect of channel estimation errors on the performances of SDCs where practical channel estimators are used.

2.2 SDC MFSK with Channel State Parameter Estimation Errors

A selection diversity combiner is often a good choice in the tradeoff between complexity and performance in wireless communication systems [36]. Similarly, non-coherent frequency shift keying (NCFSK) is used in many practical systems [37], [38]. In [39] and [40], analyses of SDC and generalized SDC using noisy channel estimates have been performed, respectively, based on the assumption of identical diversity branches and balanced noise powers. In this work, we evaluate the effect of channel estimation errors on the performance of SDC MFSK, under the assumption that all diversity branches are slowly and flatly Rayleigh faded and independent but are not identically distributed and that the diversity branch noise powers are different. This is the case, for example, when the mean powers of the fading on different diversity branches may differ because of shadowing. In space diversity, the noise powers on the branches may be different due to the different loads of the antennas [41] or some antenna processing [42]. Also, in diversity systems with interference, and when a Gaussian approximation of the interference is used, the powers of the equivalent noise terms on the branches (including both interference and noise) may be different [43]. Finally, component tolerances and aging in electronic systems result in imbalance between branches. Both maximum signal amplitude based selection and maximum SNR based selection are examined. The channel metrics are calculated by using noisy estimates of the signal amplitude and the noise power obtained from a ML estimator developed in Section 2.1, a MMSE estimator [44] or a Gans' fading correlation model [45], [39].

2.2.1 System Model

We consider the case when L_d independent diversity branches are available. Each is slowly and flatly fading. It is assumed that all signals in the MFSK signaling set have equal symbol powers, which have been normalized to equal 1. Without loss of generality, assume that the k-th signal is transmitted. Then, the output of the i-th matched filter on the l-th branch can be expressed as

$$y_{li} = A_l s_i + n_{li} \tag{2.52}$$

where A_l is the complex fading gain on the l-th branch, s_i is a constant with $s_i = 1$ for $i = k$ and $s_i = 0$ for $i = 1, 2, \cdots, M$ and $i \neq k$, and n_{li} is the complex noise component in the output of the i-th matched filter on the l-th branch. In a Rayleigh fading channel, A_l is a complex Gaussian random variable with mean zero and variance $2\alpha_l^2$. The noise components n_{li} for a fixed l and $i = 1, 2, \cdots, M$ are independent complex Gaussian random variables each with mean zero and variance N_l. Assuming that the noise processes are independent of the fading processes, the random variables n_{li}, $i = 1, 2, \cdots, M$, are independent of A_l.

If perfect knowledge of A_l and N_l is available in the SDC, one can calculate the branch SNR according to $\rho_l = \frac{|A_l|^2}{N_l}$, compare ρ_l $(l = 1, 2, \cdots, L_d)$ and select the branch corresponding to the largest SNR. Equivalently, if $N_l = N$ for $l = 1, 2, \cdots, L_d$, it suffices to calculate the signal amplitude by using $r_l = |A_l|$, compare r_l $(l = 1, 2, \cdots, L_d)$ and select the branch corresponding to the largest signal amplitude. Unfortunately, perfect knowledge of A_l and N_l is not available in practical communication systems. Instead, one has to use channel parameter estimators to obtain estimates of A_l and N_l, $\hat{A}_l$ and $\hat{N}_l$, and then use $\hat{A}_l$ and $\hat{N}_l$ to calculate the estimated branch SNR $\hat{\rho}_l = \frac{|\hat{A}_l|^2}{\hat{N}_l}$ or the estimated signal amplitude $\hat{r}_l = |\hat{A}_l|$, and finally select the best diversity branch according to $\hat{\rho}_l$ or $\hat{r}_l$. Some general assumptions on $\hat{A}_l$ and $\hat{N}_l$ ($l = 1, 2, \cdots, L_d$) are given as follows.

We assume that the channel estimate, $\hat{A}_l$, is a complex Gaussian random variable with mean zero. Denote $\hat{A}_l = \hat{A}_l^R + j\hat{A}_l^I$ and $A_l = A_l^R + jA_l^I$. Similar to [46, eq. (8-101)], it is assumed that $\hat{A}_l^R$, $\hat{A}_l^I$, A_l^R and A_l^I are jointly Gaussian distributed with

$$f(\hat{A}_l^R, \hat{A}_l^I, A_l^R, A_l^I) = \frac{1}{4\pi^2 |\mathbf{M}_l|^{\frac{1}{2}}} e^{-\frac{\alpha_l^2\left[(\hat{A}_l^R)^2+(\hat{A}_l^I)^2\right]+\tilde{\alpha}_l^2\left[(A_l^R)^2+(A_l^I)^2\right]-2R_{cl}\left[\hat{A}_l^R A_l^R+\hat{A}_l^I A_l^I\right]-2R_{sl}\left[\hat{A}_l^R A_l^I-\hat{A}_l^I A_l^R\right]}{2|\mathbf{M}_l|^{\frac{1}{2}}}} \tag{2.53}$$

where the covariance matrix is $\mathbf{M}_l = \begin{bmatrix} \tilde{\alpha}_l^2 & 0 & R_{cl} & R_{sl} \\ 0 & \tilde{\alpha}_l^2 & -R_{sl} & R_{cl} \\ R_{cl} & -R_{sl} & \alpha_l^2 & 0 \\ R_{sl} & R_{cl} & 0 & \alpha_l^2 \end{bmatrix}$ [46, eq. (8-98)]. The assumption of (2.53) includes the MMSE channel estimate derived in [44] and the ML channel estimate derived in Section 2.1 as special cases. A detailed derivation of (2.53) and the definitions of R_{cl} and R_{sl} are given in [46].

We also assume that the noise power estimate, $\hat{N}_l$, is obtained using the ML estimator developed in Section 2.1. It was derived in Section 2.1 that the ML estimate, $\hat{N}_l$, has PDF

$$f(\hat{N}_l) = \frac{1}{\Gamma(p)(2\beta_l^2)^p} \hat{N}_l^{p-1} e^{-\frac{\hat{N}_l}{2\beta_l^2}}, \quad \hat{N}_l > 0 \tag{2.54}$$

where $p = I - 1$, $\beta_l^2 = \frac{N_l}{2(p+1)}$, I is the number of available independent samples defined in Section 2.1, and $\Gamma(\cdot)$ is the Gamma function [2, p. 255]. This is, for example, the case when the noise power is estimated using data-aided estimators. This is also the case when the noise power is estimated using decision-based ML estimators in Section 2.1 and no decision errors occur. Since the noise power normally changes little during the transmission of the desired user's data in most practical wireless communication channels, it can be estimated off-line by using training sequences (which, for example, may be for synchronizer training or equalizer training). Therefore, sufficiently large p can be achieved by sampling enough symbols in the training sequences. We assume that $\hat{N}_l$ is independent of $\hat{A}_l$. This is also proved in Appendix B when $\hat{A}_l$ is obtained by using the ML estimator in Section 2.1.

Most previous works on SDC base the analysis on perfect knowledge of A_l and N_l. In practical systems, and in this work, we assume that the data recovery decision is based on the estimated quantities, $\hat{A}_l$ and $\hat{N}_l$. Assuming that the k-th signal has been sent and the l-th diversity branch has been chosen, the non-coherent data decision is made by comparing

$$z_{li} = |y_{li}| = |A_l s_i + n_{li}| \tag{2.55}$$

where $i = 1, 2, \cdots, M$. An error will occur if z_{lk} is less than any of z_{li}, $i = 1, 2, \cdots, M$ and $i \neq k$. In the sequel, we derive the average symbol error rate (ASER) of SDC using the maximum estimated signal amplitude selection criterion as well as the maximum estimated SNR criterion. For convenience, we denote the SDC using the maximum estimated signal amplitude criterion as $\hat{r}$-SDC and the SDC using the maximum estimated SNR criterion as $\hat{\rho}$-SDC.

2.2.2 $\hat{r}$-SDC

Assume that all M possible transmitted signals have equal *a priori* probabilities. The ASER of the system can be expressed as

$$\bar{P}_e = \frac{1}{M} \sum_{k=1}^{M} Pr(\hat{s} \neq s_k | s_k) \tag{2.56}$$

where $Pr(\hat{s} \neq s_k | s_k)$ is the probability that the receiver decision, $\hat{s}$, is incorrect when the k-th signal, s_k, is transmitted. Note that $\bar{P}_e$ will depend on the branch selection criterion used in SDC. If the $\hat{r}$-SDC criterion is used, the combiner chooses the branch with the largest estimated signal amplitude and uses matched filter outputs on that branch to make the data decision. Denote $\hat{r}_l = |\hat{A}_l|$ and $r_l = |A_l|$ as the estimate and the true value of the signal amplitude on the l-th branch, respectively. Then, the l-th branch is chosen when $\hat{r}_l^2 \geq \hat{r}_j^2$, $j = 1, 2, \cdots, L_d, j \neq l$, and the error rate in (2.56) can be rewritten as

$$\bar{P}_e = \frac{1}{M} \sum_{k=1}^{M} \sum_{l=1}^{L_d} Pr(\hat{s}_l \neq s_k, \hat{r}_l^2 \geq \hat{r}_j^2, j = 1, 2, \cdots, L_d, j \neq l | s_k) \tag{2.57}$$

where $\hat{s}_l$ denotes the data decision made by using matched filter outputs on the l-th branch. Conditioning the error rate in (2.57) on $\hat{r}_l$ and r_l gives

$$\bar{P}_e = \frac{1}{M} \sum_{k=1}^{M} \sum_{l=1}^{L_d} \int_0^{\infty} \int_0^{\infty} Pr(\hat{s}_l \neq s_k, \hat{r}_l^2 \geq \hat{r}_j^2, j = 1, 2, \cdots, L_d, j \neq l | \hat{r}_l, r_l, s_k) \cdot f(\hat{r}_l, r_l) d\hat{r}_l dr_l. \tag{2.58}$$

One sees from (2.58) that $Pr(\hat{s}_l \neq s_k, \hat{r}_l^2 \geq \hat{r}_j^2, j = 1,2,\cdots,L_d, j \neq l | \hat{r}_l, r_l, s_k)$ is the joint conditional probability of L_d statistical events, $\hat{s}_l \neq s_k$ and $\hat{r}_l^2 \geq \hat{r}_j^2$ ($j = 1,2,\cdots,L_d$ and $j \neq l$). Given $\hat{r}_l$ and r_l, the statistical event $\hat{s}_l \neq s_k$ is determined solely by the random variables n_{li}, $i = 1,2,\cdots,M$, while the statistical events $\hat{r}_l^2 \geq \hat{r}_j^2$ ($j = 1,2,\cdots,L_d$ and $j \neq l$) are determined solely by the random variables $\hat{r}_j$ ($j = 1,2,\cdots,L_d$ and $j \neq l$). Since n_{li} ($i = 1,2,\cdots,M$) and $\hat{r}_j$ ($j = 1,2,\cdots,L_d$ and $j \neq l$) are independent, it is clear that

$$\bar{P}_e = \frac{1}{M} \sum_{k=1}^{M} \sum_{l=1}^{L_d} \int_0^\infty \int_0^\infty Pr(\hat{s}_l \neq s_k | r_l, s_k) \cdot \left[\prod_{j=1, j\neq l}^{L_d} Pr(\hat{r}_l^2 \geq \hat{r}_j^2 | \hat{r}_l) \right] \cdot f(\hat{r}_l, r_l) d\hat{r}_l dr_l. \tag{2.59}$$

The conditional probability $Pr(\hat{s}_l \neq s_k | r_l, s_k)$ is actually the ASER of the NCFSK signals in a static AWGN channel. It is well known that [1]

$$Pr(\hat{s}_l \neq s_k | r_l, s_k) = \sum_{m=1}^{M-1} (-1)^{m+1} \binom{M-1}{m} \frac{1}{m+1} \exp\left\{ -\frac{m r_l^2}{(m+1) N_l} \right\}. \tag{2.60}$$

The conditional probability $Pr(\hat{r}_l^2 \geq \hat{r}_j^2 | \hat{r}_l)$ is the value of the cumulative density function (CDF) of $\hat{r}_j^2$ at $\hat{r}_l^2$, that is, $Pr(\hat{r}_l^2 \geq \hat{r}_j^2 | \hat{r}_l) = F_{\hat{r}_j^2}(\hat{r}_l^2)$. Since $\hat{A}_j$ is a complex Gaussian random variable with mean zero and variance $2\tilde{\alpha}_j^2$, $\hat{r}_j^2 = |\hat{A}_j|^2$ has CDF

$$F_{\hat{r}_j^2}(x) = 1 - e^{-\frac{x}{2\tilde{\alpha}_j^2}}, \quad x > 0 \tag{2.61}$$

for each $j = 1,2,\cdots,L_d$ and $j \neq l$. From (2.61), one has

$$Pr(\hat{r}_l^2 \geq \hat{r}_j^2 | \hat{r}_l) = 1 - e^{-\frac{\hat{r}_l^2}{2\tilde{\alpha}_j^2}} \tag{2.62}$$

and

$$\prod_{j=1, j\neq l}^{L_d} Pr(\hat{r}_l^2 \geq \hat{r}_j^2 | \hat{r}_l) = \sum_{n_1=0}^{L_d-1} \sum_{n_2=1}^{Q} (-1)^{n_1} e^{-\eta \hat{r}_l^2} \tag{2.63}$$

where $Q = \binom{L_d-1}{n_1}$ and $\eta = \sum_{k=1}^{n_1} \frac{1}{2\tilde{\alpha}_{k_{n_2}}^2}$ ($\eta = 0$ when $n_1 = 0$). The index k_{n_2} denotes the k-th number in the n_2-th combination obtained by selecting n_1 numbers out of $L_d - 1$ numbers $j = 1,2,\cdots,L_d$ and $j \neq l$. For example, when $L_d = 3$, n_1 can be 0, 1, or 2. When $n_1 = 1$ and $l = 1$, $\sum_{n_2=1}^{Q} e^{-\eta \hat{r}_l^2} = e^{-\frac{\hat{r}_l^2}{2\tilde{\alpha}_2^2}} + e^{-\frac{\hat{r}_l^2}{2\tilde{\alpha}_3^2}}$ since $k_1 = 2$ and $k_2 = 3$ in this case.

The joint PDF of $\hat{r}_l$ and r_l is now derived as follows. Denote $\hat{\theta}_l = \arctan(\hat{A}_l^I, \hat{A}_l^R)$ and $\theta_l = \arctan(A_l^I, A_l^R)$ as the phase angles of $\hat{A}_l$ and A_l, respectively. Also, we have $\hat{r}_l = \sqrt{(\hat{A}_l^R)^2 + (\hat{A}_l^I)^2}$

and $r_l = \sqrt{(A_l^R)^2 + (A_l^I)^2}$. Following similar procedures to those in [46, p. 163], it is derived from (2.53) that the joint PDF of $\hat{r}_l$, r_l, $\hat{\theta}_l$ and θ_l is

$$f(\hat{r}_l, r_l, \hat{\theta}_l, \theta_l) = \frac{\hat{r}_l r_l}{4\pi^2 \Delta_l} \cdot e^{-\frac{1}{2\Delta_l}\left[\alpha_l^2 \hat{r}_l^2 + \tilde{\alpha}_l^2 r_l^2 - 2R_{cl} r_l \hat{r}_l \cos(\theta_l - \hat{\theta}_l) - 2R_{sl} r_l \hat{r}_l \sin(\theta_l - \hat{\theta}_l)\right]},$$
$$\hat{r}_l > 0, r_l > 0, 2\pi > \hat{\theta}_l, \theta_l > 0 \tag{2.64}$$

where $\Delta_l = |\mathbf{M}_l|^{\frac{1}{2}} = \tilde{\alpha}_l^2 \alpha_l^2 - R_{cl}^2 - R_{sl}^2$ is the square root of the determinant of $\mathbf{M}_l$. By integrating $\hat{\theta}_l$ and θ_l from 0 to 2π in (2.64), the joint PDF of r_l and $\hat{r}_l$ can be derived as

$$f(\hat{r}_l, r_l) = \frac{\hat{r}_l r_l}{\Delta_l} e^{-\frac{\alpha_l^2}{2\Delta_l}\hat{r}_l^2 - \frac{\tilde{\alpha}_l^2}{2\Delta_l} r_l^2} I_0\left(\frac{\hat{r}_l r_l}{\Delta_l}\sqrt{R_{cl}^2 + R_{sl}^2}\right), \quad \hat{r}_l > 0, r_l > 0 \tag{2.65}$$

where $I_0(\cdot)$ is the zero-th order modified Bessel function of the first kind [2, p. 374]. Note that (2.65) agrees with [46, eqn. (8-103)] when $\alpha_l^2 = \tilde{\alpha}_l^2$.

Using (2.60), (2.63) and (2.65), the integral in (2.59) can be solved. This is done as follows. Denote

$$U_l(v) = \int_0^\infty \int_0^\infty e^{-v r_l^2} \left[\prod_{j=1, j\neq l}^{L} Pr(\hat{r}_l^2 \geq \hat{r}_j^2 | \hat{r}_l)\right] \cdot f(\hat{r}_l, r_l) d\hat{r}_l dr_l. \tag{2.66}$$

Inserting (2.60) in (2.59), one has

$$\bar{P}_e = \sum_{l=1}^{L_d} \sum_{m=1}^{M-1} (-1)^{m+1} \binom{M-1}{m} \frac{1}{m+1} U_l\left(\frac{m}{(m+1)N_l}\right). \tag{2.67}$$

The integration in (2.66) is solved to give

$$U_l(v) = \sum_{n_1=0}^{L_d-1} \sum_{n_2=1}^{Q} \frac{(-1)^{n_1}}{2v\alpha_l^2 + \eta(4v\Delta_l + 2\tilde{\alpha}_l^2) + 1} \tag{2.68}$$

where [35, eqn. (6.631.1)] is used. Finally, the ASER of $\hat{r}$-SDC is obtained by putting (2.68) into (2.67), giving

$$\bar{P}_e = \sum_{l=1}^{L_d} \sum_{m=1}^{M-1} \sum_{n_1=0}^{L_d-1} \sum_{n_2=1}^{N} \frac{(-1)^{m+n_1+1} \binom{M-1}{m}}{m\bar{\rho}_l + 2\tilde{\alpha}_l^2 \eta (m\bar{\rho}_l \frac{\Delta_l}{\tilde{\alpha}_l^2 \alpha_l^2} + m + 1) + m + 1} \tag{2.69}$$

where $\bar{\rho}_l = \frac{2\alpha_l^2}{N_l}$ is the average SNR per symbol on the l-th branch. Defining $c_l = \frac{E\{\hat{A}_l A_l^*\}}{\sqrt{Var\{\hat{A}_l\} Var\{A_l\}}}$ as the correlation coefficient between $\hat{A}_l$ and A_l, it can be derived that $|c_l|^2 = \frac{R_{cl}^2 + R_{sl}^2}{\tilde{\alpha}_l^2 \alpha_l^2}$. Rewriting (2.69) using c_l, one has

$$\bar{P}_e = \sum_{l=1}^{L_d} \sum_{m=1}^{M-1} \sum_{n_1=0}^{L_d-1} \sum_{n_2=1}^{Q} \frac{(-1)^{m+n_1+1} \binom{M-1}{m}}{m\bar{\rho}_l + 2\tilde{\alpha}_l^2 \eta \left[m\bar{\rho}_l (1 - |c_l|^2) + m + 1\right] + m + 1}, \tag{2.70}$$

which is the ASER of $\hat{r}$-SDC with channel estimation errors. Note that the only estimation error in $\hat{r}$-SDC occurs in the estimation of the channel state parameter, the signal amplitude, which is completely described by c_l in (2.70). The result in (2.70) is general; for example, the result for balanced noise powers can be obtained from (2.70) by letting $N_l = N$ for $l = 1, 2, \cdots, L_d$. Furthermore, (2.70) applies to various channel state parameter estimator scenarios. Next, we discuss some special cases.

First, if the ML estimator in Section 2.1 is used to estimate the branch signal amplitude, when there are no decision errors in the estimation, one has from (2.36)

$$\hat{A}_l = A_l + g_l \tag{2.71}$$

where g_l is a complex Gaussian random error independent of a_l, and has mean zero and variance $\frac{N_l}{q}$ with $q = I$ being the number of independent samples used in the estimation as defined in Section 2.1. Note that the signal power has been normalized equal to 1, as mentioned previously. From (2.71), it can be derived that

$$\tilde{\alpha}_l^2 = \alpha_l^2 + \frac{N_l}{2q}, \quad R_{cl} = \alpha_l^2, \quad R_{sl} = 0 \tag{2.72}$$

for $l = 1, 2, \cdots, L_d$. The ASER of $\hat{r}$-SDC in this case is obtained by putting (2.72) into (2.70), giving

$$\bar{P}_e = \sum_{l=1}^{L_d} \sum_{m=1}^{M-1} \sum_{n_1=0}^{L_d-1} \sum_{n_2=1}^{Q} \frac{(-1)^{m+n_1+1} \binom{M-1}{m}}{m\bar{\rho}_l + b\left[\frac{m\bar{\rho}_l}{q\bar{\rho}_l+1} + m + 1\right] + m + 1} \tag{2.73}$$

where $b = \sum_{k=1}^{n_1} \frac{q\alpha_l^2 + N_l/2}{q\alpha_{k_{n_2}}^2 + N_{k_{n_2}}/2}$ ($b = 0$ when $n_1 = 0$). Assuming identical diversity branches and equal noise powers, the branch parameters satisfy

$$\tilde{\alpha}_l^2 = \alpha^2 + \frac{N}{2q}, \quad \alpha_l^2 = \alpha^2, \quad R_{cl} = \alpha^2, \quad R_{sl} = 0, \quad N_l = N \tag{2.74}$$

where $l = 1, 2, \cdots, L_d$. Using (2.74) in (2.73), it can be shown that

$$\bar{P}_e = \sum_{m=1}^{M-1} \sum_{n_1=0}^{L_d-1} \frac{(-1)^{m+n_1+1} \binom{M-1}{m} L_d Q}{m\bar{\rho} + n_1\left[\frac{m\bar{\rho}}{q\bar{\rho}+1} + m + 1\right] + m + 1}. \tag{2.75}$$

One sees from (2.73) and (2.75) that the ASER of $\hat{r}$-SDC in this case is related to q. This is expected, as q determines the estimation error in the estimation of A_l, which is the only estimation error in $\hat{r}$-SDC. The results in (2.73) and (2.75) are also valid when data-aided estimators are used.

Second, if the MMSE estimator is used to estimate the signal amplitude, one has [44]

$$\hat{A}_l = A_l - g_l \tag{2.76}$$

where g_l is independent of $\hat{A}_l$, and is a complex Gaussian random error with mean zero and variance $2\alpha_{gl}^2$. It is then derived from (2.76) that

$$\alpha_l^2 = \tilde{\alpha}_l^2 + \alpha_{gl}^2, \quad R_{cl} = \tilde{\alpha}_l^2, \quad R_{sl} = 0, \tag{2.77}$$

for $l = 1, 2, \cdots, L_d$. Using (2.77) in (2.70), the ASER of $\hat{r}$-SDC when the MMSE estimator is used to obtain $\hat{A}_l$ can be derived as

$$\bar{P}_e = \sum_{l=1}^{L_d} \sum_{m=1}^{M-1} \sum_{n_1=0}^{L_d-1} \sum_{n_2=1}^{Q} \frac{(-1)^{m+n_1+1} \binom{M-1}{m}}{m\bar{\rho}_l + 2\tilde{\alpha}_l^2 \eta \left[m\bar{\rho}_l (1 - \frac{\tilde{\alpha}_l^2}{\alpha_l^2}) + m + 1 \right] + m + 1}. \tag{2.78}$$

When all the diversity branches are identically distributed and the noise powers on all diversity branches are the same, one also has from (2.77)

$$\tilde{\alpha}_l^2 = \tilde{\alpha}^2, \quad \alpha_l^2 = \alpha^2 = \tilde{\alpha}^2 + \alpha_g^2, \quad R_{cl} = \tilde{\alpha}^2, \quad R_{sl} = 0, \quad N_l = N \tag{2.79}$$

where $l = 1, 2, \cdots, L_d$. Similarly, by using (2.79) in (2.78), it can be derived that

$$\bar{P}_e = \sum_{m=1}^{M-1} \sum_{n_1=0}^{L_d-1} \frac{(-1)^{m+n_1+1} \binom{M-1}{m} L_d Q}{m\bar{\rho} + n_1 \left[m\bar{\rho} (1 - \frac{\tilde{\alpha}^2}{\alpha^2}) + m + 1 \right] + m + 1}. \tag{2.80}$$

Fianally, in Gans' fading correlation model [45], [39], an additional assumption of

$$\alpha_l^2 = \tilde{\alpha}_l^2 \tag{2.81}$$

for $l = 1, 2, \cdots, L_d$ is used. This applies to the case where estimation error is due solely to temporal decorrelation between the estimate and the processed data symbol and not to independent estimator noise [47], [48]. Putting (2.81) into (2.70), the ASER of $\hat{r}$-SDC becomes

$$\bar{P}_e = \sum_{l=1}^{L_d} \sum_{m=1}^{M-1} \sum_{n_1=0}^{L_d-1} \sum_{n_2=1}^{Q} \frac{(-1)^{m+n_1+1} \binom{M-1}{m}}{m\bar{\rho}_l + 2\alpha_l^2 \eta' \left[m\bar{\rho}_l (1 - |c_l|^2) + m + 1 \right] + m + 1} \tag{2.82}$$

where $\eta' = \sum_{k=1}^{n_1} \frac{1}{2\alpha_{k_{n_2}}^2}$ ($\eta' = 0$ when $n_1 = 0$) and $|c_l|^2 = \frac{R_{cl}^2 + R_{sl}^2}{\alpha_l^4}$. If the fading on each branch is identically distributed and the noise power is the same, one further has

$$\tilde{\alpha}_l^2 = \alpha_l^2 = \alpha^2, \quad R_{cl} = R_c, \quad R_{sl} = R_s, \quad N_l = N \tag{2.83}$$

where $l = 1, 2, \cdots, L_d$. Using (2.83) in (2.82), the ASER of $\hat{r}$-SDC can be simplified as

$$\bar{P}_e = \sum_{m=1}^{M-1} \sum_{n_1=0}^{L_d-1} \frac{(-1)^{m+n_1+1}\binom{M-1}{m} L_d Q}{m\bar{\rho} + n_1\left[m\bar{\rho}(1-|c_l|^2) + m + 1\right] + m + 1} \tag{2.84}$$

It is easy to verify that (2.84) agrees with [39, eqn. (25)] for NCFSK signaling. Therefore, our result in (2.70) includes the corresponding result in [39] as a special case.

As the last part of this section, we derive the ASER of SDC when perfect knowledge of A_l is available. Denote it as r-SDC. The performance of r-SDC serves as a benchmark for (2.70). In the case when A_l is known, (2.56) becomes

$$\bar{P}_e = \frac{1}{M} \sum_{k=1}^{M} \sum_{l=1}^{L_d} \int_0^\infty Pr(\hat{s}_l \neq s_k, r_l^2 \geq r_j^2, j = 1, 2, \cdots, L_d, j \neq l | r_l, s_k) f(r_l) dr_l. \tag{2.85}$$

Note that the CDF of r_j^2 is

$$F_{r_j^2}(x) = 1 - e^{-\frac{x}{2\alpha_j^2}}, \quad x > 0. \tag{2.86}$$

Employing similar techniques, it can be derived that

$$\bar{P}_e = \sum_{l=1}^{L_d} \sum_{m=1}^{M-1} \sum_{n_1=0}^{L_d-1} \sum_{n_2=1}^{Q} \frac{(-1)^{m+n_1+1}\binom{M-1}{m}}{m\bar{\rho}_l + (1 + 2\alpha_l^2\eta')(m+1)} \tag{2.87}$$

where η' is defined as before. One sees that (2.87) can also be obtained from (2.70) by letting $|c_l| = 1$ and $\tilde{\alpha}_l^2 = \alpha_l^2$ in (2.70). This is expected as $\hat{r}$-SDC without estimation errors will have $\hat{A}_l = A_l$, which in turn gives $|c_l| = 1$ and $\tilde{\alpha}_l^2 = \alpha_l^2$. Next, we derive the ASER of $\hat{\rho}$-SDC.

2.2.3 $\hat{\rho}$-SDC

Using similar reasoning as previously, the ASER of $\hat{\rho}$-SDC can also be obtained. In $\hat{\rho}$-SDC, the branch with the largest estimated SNR is selected and the error rate (2.56) is expressed as

$$\bar{P}_e = \frac{1}{M} \sum_{k=1}^{M} \sum_{l=1}^{L_d} Pr(\hat{s}_l \neq s_k, \hat{\rho}_l \geq \hat{\rho}_j, j = 1, 2, \cdots, L_d, j \neq l | s_k) \tag{2.88}$$

where $\hat{\rho}_l = \frac{\hat{r}_l^2}{\hat{N}_l}$ is the estimated SNR on the l-th branch. Similarly, by using the conditional probability and the independence of the random variables n_{li} ($i = 1, 2, \cdots, M$) and $\hat{\rho}_j$ ($j = 1, 2, \cdots, L_d$ and $j \neq l$), one has from (2.88)

$$\bar{P}_e = \frac{1}{M} \sum_{k=1}^{M} \sum_{l=1}^{L_d} \int_0^\infty \int_0^\infty Pr(\hat{s}_l \neq s_k | r_l, s_k) \cdot \left[\prod_{j=1, j\neq l}^{L_d} Pr(\hat{\rho}_l \geq \hat{\rho}_j | \hat{\rho}_l)\right] \cdot f(r_l, \hat{\rho}_l) dr_l d\hat{\rho}_l \tag{2.89}$$

where $Pr(\hat{s}_l \neq s_k | r_l, s_k)$ is given by (2.60).

The conditional probability $Pr(\hat{\rho}_l \geq \hat{\rho}_j | \hat{\rho}_l)$ is derived as follows. Since $\hat{r}_j$ is a Rayleigh random variable, $\hat{r}_j^2$ has a central chi-square distribution with two degrees of freedom (an exponential distribution) and parameter $\tilde{\alpha}_j^2$ [1]. Also, since $\hat{N}_j$ has a central chi-square distribution with $2p$ degrees of freedom and parameter β_j^2 as given in (2.54), the random variable $\hat{\rho}_j = \frac{\hat{r}_j^2}{\hat{N}_j}$ has PDF [49, eqn. (27.3)]

$$f_{\hat{\rho}_j}(x) = \frac{p\beta_j^2}{\tilde{\alpha}_j^2}\left(1 + \frac{\beta_j^2}{\tilde{\alpha}_j^2}x\right)^{-(p+1)}, \quad x > 0. \tag{2.90}$$

The CDF of $\hat{\rho}_j$ is then derived from (2.90) by integration as

$$F_{\hat{\rho}_j}(x) = 1 - \frac{1}{\left(1 + \frac{\beta_j^2}{\tilde{\alpha}_j^2}x\right)^p}, \quad x > 0. \tag{2.91}$$

Using (2.91), one finally has

$$Pr(\hat{\rho}_l \geq \hat{\rho}_j | \hat{\rho}_l) = 1 - \frac{1}{\left(1 + \frac{\beta_j^2}{\tilde{\alpha}_j^2}\hat{\rho}_l\right)^p}. \tag{2.92}$$

The joint PDF of r_l and $\hat{\rho}_l$ can be derived as follows. Since $\hat{N}_l$ is independent of $\hat{A}_l$ and A_l, one has

$$f(\hat{r}_l, r_l, \hat{N}_l) = f(\hat{r}_l, r_l) f(\hat{N}_l) \tag{2.93}$$

where $f(\hat{r}_l, r_l)$ and $f(\hat{N}_l)$ are given by (2.65) and (2.54), respectively. By replacing $\hat{N}_l$ with $\hat{\rho}_l$ in (2.93) and integrating the result with respect to $\hat{r}_l$, one has

$$f(r_l, \hat{\rho}_l) = \frac{p\beta_l^2}{\Delta_l} \cdot \frac{r_l}{\left(1 + \frac{\beta_l^2\alpha_l^2}{\Delta_l}\hat{\rho}_l\right)^{p+1}} e^{-\frac{\alpha_l^2}{2\Delta_l}r_l^2} \cdot {}_1F_1\left(p+1, 1; \frac{r_l^2(R_{cl}^2 + R_{sl}^2)}{2\Delta_l\alpha_l^2 + \frac{2\Delta_l^2}{\beta_l^2\hat{\rho}_l}}\right), \quad r_l > 0, \hat{\rho}_l > 0 \tag{2.94}$$

where ${}_1F_1(\cdot, \cdot; \cdot)$ is the confluent hypergeometric function [2, p. 504].

Using (2.60), (2.92) and (2.94) in (2.89), the ASER of $\hat{\rho}$-SDC can be derived. Again, denoting

$$W_l(v) = \int_0^\infty \int_0^\infty e^{-vr_l^2} \left[\prod_{j=1, j\neq l}^{L_d} Pr(\hat{\rho}_l > \hat{\rho}_j | \hat{\rho}_l)\right] \cdot f(r_l, \hat{\rho}_l) dr_l d\hat{\rho}_l, \tag{2.95}$$

(2.89) can be written in a form similar to (2.67), except that $U_l(v)$ is replaced by $W_l(v)$ now. Using (2.92) and (2.94), together with [2, eqn. (7.621.5)], it can be shown that

$$W_l(v) = \frac{p\beta_l^2}{2\Delta_l v + \tilde{\alpha}_l^2} \int_0^{\infty} \prod_{j=1, j\neq l}^{L_d} \left[1 - \frac{1}{(1 + \frac{\beta_j^2}{\tilde{\alpha}_j^2}\hat{\rho}_l)^p} \right] \cdot \frac{1}{(1 + \frac{2v\alpha_l^2 + 1}{2v\Delta_l + \tilde{\alpha}_l^2} \beta_l^2 \hat{\rho}_l)^{p+1}} d\hat{\rho}_l. \tag{2.96}$$

Finally, the ASER of the $\hat{\rho}$-SDC is derived from (2.96) as

$$\begin{aligned} \bar{P}_e &= \sum_{l=1}^{L_d} \sum_{m=1}^{M-1} \frac{(-1)^{m+1} \binom{M-1}{m} p}{m\bar{\rho}_l(1 - |c_l|^2) + m + 1} \cdot \frac{\beta_l^2}{\tilde{\alpha}_l^2} \\ &\quad \cdot \int_0^{\infty} \prod_{j=1, j\neq l}^{L_d} \left[1 - \frac{1}{(1 + \frac{\beta_j^2}{\bar{\sigma}_j^2}\hat{\rho}_l)^p} \right] \cdot \frac{1}{(1 + \frac{m\bar{\rho}_l + m + 1}{m\bar{\rho}_l(1 - |c_l|^2) + m + 1} \frac{\beta_l^2}{\tilde{\alpha}_l^2} \hat{\rho}_l)^{p+1}} d\hat{\rho}_l \end{aligned} \tag{2.97}$$

which can be evaluated numerically. Note that, in $\hat{\rho}$-SDC, estimation error not only occurs in the estimation of the branch signal amplitude but also occurs in the estimation of the branch noise power. These two components are completely described by c_l and p. Note further that (2.97) will specialize to (2.70) when $p \to \infty$ and $N_l = N$ ($l = 1, 2, \cdots, L$). This can be verified by putting $\beta_l^2 = \frac{N_l}{2(p+1)}$ and $N_l = N$ ($l = 1, 2, \cdots, L$) into (2.97), using $\lim_{p\to\infty} \left(1 + \frac{x}{p}\right)^p = e^x$ [50, eqn. (2.9.2.7)], and solving the resulting integral. Intuitively, this is also expected as $\hat{r}$-SDC is a special case of $\hat{\rho}$-SDC when noise powers on all branches are the same and the noise power estimation in the $\hat{\rho}$-SDC is free of errors. Moreover, a closed-form expression of (2.97) for $L_d = 2$ can be derived as

$$\begin{aligned} \bar{P}_e &= \sum_{l,j=1, l\neq j}^{2} \sum_{m=1}^{M-1} \frac{(-1)^{m+1} \binom{M-1}{m}}{m\bar{\rho}_l + m + 1} \cdot \left[1 - \frac{\tilde{\alpha}_j^2}{2\beta_j^2} \cdot \frac{\beta_l^2}{\tilde{\alpha}_l^2} \cdot \frac{m\bar{\rho}_l + m + 1}{m\bar{\rho}_l(1 - |c_l|^2) + m + 1} \right. \\ &\quad \left. \cdot F\left(p + 1, 1; 2p + 1; 1 - \frac{m\bar{\rho}_l + m + 1}{m\bar{\rho}_l(1 - |c_l|^2) + m + 1} \cdot \frac{\beta_l^2}{\tilde{\alpha}_l^2} \cdot \frac{\tilde{\alpha}_j^2}{\beta_j^2} \right) \right] \end{aligned} \tag{2.98}$$

where $F(\cdot, \cdot; \cdot; \cdot)$ is the hypergeometric function [2, p. 556]. As previously, results for three special cases can be obtained from (2.97), given next.

The ASER of $\hat{\rho}$-SDC when the ML estimator is used to estimate the channel state parameter, the signal amplitude, can be derived as follows. Putting (2.72) into (2.97), (2.97) becomes

$$\bar{P}_e = \sum_{l=1}^{L_d} \sum_{m=1}^{M-1} \frac{(-1)^{m+1} \binom{M-1}{m}}{\frac{m\bar{\rho}_l}{1+q\bar{\rho}_l} + m + 1} \cdot \frac{pq}{q(p+1)\bar{\rho}_l + p + 1} \tag{2.99}$$

$$\cdot \int_0^{\infty} \prod_{j=1, j\neq l}^{L_d} \left[1 - \frac{1}{(1+\frac{q}{q(p+1)\bar{\rho}_j+p+1}\hat{\rho}_l)^p}\right] \cdot \frac{1}{(1+\frac{m\bar{\rho}_l+m+1}{\frac{m\bar{\rho}_l}{1+q\bar{\rho}_l}+m+1}\frac{q}{q(p+1)\bar{\rho}_l+p+1}\hat{\rho}_l)^{p+1}} d\hat{\rho}_l.$$

By putting (2.74) in (2.99), one further has

$$\bar{P}_e = \sum_{m=1}^{M-1}\sum_{n_1=0}^{L_d-1} \frac{(-1)^{m+n_1+1}\binom{M-1}{m}\binom{L_d-1}{n_1}L_d}{n_1+1} \cdot \frac{F\left(p+1,1;p(n_1+1)+1;-\frac{m\bar{\rho}}{\frac{m\bar{\rho}}{1+q\bar{\rho}}+m+1}\frac{q\bar{\rho}}{q\bar{\rho}+1}\right)}{\frac{m\bar{\rho}}{1+q\bar{\rho}}+m+1} \tag{2.100}$$

which is the ASER of $\hat{\rho}$-SDC for identical branches. Comparing (2.99) and (2.100) with (2.73) and (2.75), one sees that the ASER of $\hat{\rho}$-SDC in this case is not only related to q but also to p, as the estimation errors in $\hat{\rho}$-SDC result from the estimation of both A_l and N_l. Again, these results are also valid for data-aided estimation.

Similarly, by putting (2.77) into (2.97), one has

$$\begin{aligned}\bar{P}_e &= \sum_{l=1}^{L_d}\sum_{m=1}^{M-1} \frac{(-1)^{m+1}\binom{M-1}{m}p}{m\bar{\rho}_l(1-\frac{\tilde{\alpha}_l^2}{\alpha_l^2})+m+1}\frac{\beta_l^2}{\tilde{\alpha}_l^2} \\ &\cdot \int_0^{\infty} \prod_{j=1, j\neq l}^{L_d} \left[1 - \frac{1}{(1+\frac{\beta_j^2}{\tilde{\alpha}_j^2}\hat{\rho}_l)^p}\right] \cdot \frac{1}{(1+\frac{m\bar{\rho}_l+m+1}{m\bar{\rho}_l(1-\frac{\tilde{\alpha}_l^2}{\alpha_l^2})+m+1}\frac{\beta_l^2}{\tilde{\alpha}_l^2}\hat{\rho}_l)^{p+1}} d\hat{\rho}_l.\end{aligned} \tag{2.101}$$

The corresponding result for identical branches can be shown to be

$$\bar{P}_e = \sum_{m=1}^{M-1}\sum_{n_1=0}^{L_d-1} \frac{(-1)^{m+n_1+1}\binom{M-1}{m}\binom{L_d-1}{n_1}L_d}{n_1+1} \cdot \frac{F\left(p+1,1;p(n_1+1)+1;-\frac{m\bar{\rho}\frac{\tilde{\alpha}^2}{\sigma^2}}{m\bar{\rho}(1-\frac{\tilde{\alpha}^2}{\alpha^2})+m+1}\right)}{m\bar{\rho}(1-\frac{\tilde{\alpha}^2}{\alpha^2})+m+1} \tag{2.102}$$

To derive the ASER of $\hat{\rho}$-SDC for Gans' correlation model, one only needs to put (2.81) into (2.97), giving

$$\begin{aligned}\bar{P}_e &= \sum_{l=1}^{L_d}\sum_{m=1}^{M-1} \frac{(-1)^{m+1}\binom{M-1}{m}p}{m\bar{\rho}_l(1-|c_l|^2)+m+1}\frac{\beta_l^2}{\alpha_l^2} \\ &\cdot \int_0^{\infty} \prod_{j=1, j\neq l}^{L_d} \left[1 - \frac{1}{(1+\frac{\beta_j^2}{\alpha_l^2}\hat{\rho}_l)^p}\right] \cdot \frac{1}{(1+\frac{m\bar{\rho}_l+m+1}{m\bar{\rho}_l(1-|c_l|^2)+m+1}\frac{\beta_l^2}{\alpha_l^2}\hat{\rho}_l)^{p+1}} d\hat{\rho}_l\end{aligned} \tag{2.103}$$

By using (2.83) in (2.103) and solving the resulting integral, the ASER of $\hat{\rho}$-SDC can be determined as

$$\bar{P}_e = \sum_{m=1}^{M-1}\sum_{n_1=0}^{L_d-1} \frac{(-1)^{m+n_1+1}\binom{M-1}{m}\binom{L_d-1}{n_1}L_d}{n_1+1} \cdot \frac{F\left(p+1,1;p(n_1+1)+1;-\frac{m\bar{\rho}|c_l|^2}{m\bar{\rho}(1-|c_l|^2)+m+1}\right)}{m\bar{\rho}(1-|c_l|^2)+m+1} \tag{2.104}$$

when all the diversity branches are identically distributed and the noise powers are equal.

Finally, we derive the ASER of SDC when perfect knowledge of $\rho_l = \frac{r_l^2}{N_l}$ is available. Denote this as ρ-SDC. This will serve as a benchmark for (2.97). Again, with true values of the branch SNRs available, one has from (2.56)

$$\bar{P}_e = \frac{1}{M}\sum_{k=1}^{M}\sum_{l=1}^{L_d}\int_0^\infty Pr(\hat{s}_l \neq s_k|\rho_l, s_k) \cdot \left[\prod_{j=1, j\neq l}^{L_d} Pr(\rho_l > \rho_j|\rho_l)\right] \cdot f(\rho_l)d\rho_l. \tag{2.105}$$

Since r_l is a Rayleigh random variable, it is derived that ρ_l has CDF

$$F_{\rho_l}(x) = 1 - e^{-\frac{N_l}{2\alpha_l^2}x}, \quad x > 0. \tag{2.106}$$

By using similar techniques and reasonings to those in the previous subsection, one has

$$\bar{P}_e = \sum_{l=1}^{L_d}\sum_{m=1}^{M-1}\sum_{n_1=0}^{L_d-1}\sum_{n_2=1}^{N} \frac{(-1)^{m+n_1+1}\binom{M-1}{m}}{m\bar{\rho}_l + (1+d)(m+1)} \tag{2.107}$$

where $d = \bar{\rho}_l \sum_{k=1}^{n_1} \frac{1}{\bar{\rho}_{k_{n_3}}} (d = 0$ when $n_1 = 0)$. As previously, (2.107) can also be obtained from (2.97) by letting $|c_l| = 1$, $\tilde{\alpha}_l^2 = \alpha_l^2$ and $p \to \infty$, as a correct estimate of A_l gives $|c_l| = 1$ and $\tilde{\alpha}_l^2 = \alpha_l^2$, and a correct estimate of N_l results as $p \to \infty$. Note further that (2.87) is a special case of (2.107) when $N_l = N$ for $l = 1, 2, \cdots, L_d$, which is also expected.

2.2.4 Numerical Examples

Here, we present some numerical examples to show the effect of channel estimation errors on the performance of SDC MFSK. Consider non-coherent binary frequency shift keying (NC-BFSK) signaling. Dual-branch diversity is used. The two branches are assumed to have identical Rayleigh distributions, but may have different noise powers. Other cases can be examined accordingly. A slowly fading channel is assumed.

Fig. 2.1 shows the performance of $\hat{r}$-SDC when different values of q are used to obtain $\hat{A}_l$. As mentioned before, the only estimation error in $\hat{r}$-SDC comes from $\hat{A}_l$, which is completely determined by q. One sees from Fig. 2.1 that there is a performance penalty of about 0.5 dB between $\hat{r}$-SDC and r-SDC at $q = 2$ when the ASER is 10^{-2}. When $q = 10$ or $q = 50$, the performance difference between $\hat{r}$-SDC and r-SDC is essentially negligible for all indicated values of SNR. Therefore, in this case, a value of $q = 10$ is a good design choice. Fig. 2.2 shows the performance of $\hat{\rho}$-SDC when different values of p are used in the estimation of the branch noise power. Unlike $\hat{r}$-SDC, two sources of estimation errors occur in $\hat{\rho}$-SDC, resulting from both $\hat{A}_l$ and $\hat{N}_l$. Since the effect of the estimation error in $\hat{A}_l$ has already been studied in Fig. 2.1, only the effect of the estimation error in $\hat{N}_l$ is examined in Fig. 2.2. To see the effect of estimation errors in $\hat{N}_l$ more clearly, we set q equal to 10 to largely eliminate the performance degradation caused by $\hat{A}_l$. From Fig. 2.2, one observes that there is a penalty of about 1.2 dB between $\hat{\rho}$-SDC and ρ-SDC for $p = 2$ when the ASER is 10^{-2}, caused mainly by the estimation error in $\hat{N}_l$. The penalty decreases as p increases. Based on the improvement seen in going from $p = 2$ to $p = 10$, to $p = 50$, it is concluded that a value of $p = 50$ can be used to eliminate most of the penalty caused by $\hat{N}_l$.

Fig. 2.3 shows the performances of $\hat{\rho}$-SDC at $p = 50, q = 10$, $p = 25, q = 5$, $p = 12, q = 2$ and $p = 6, q = 1$. The values of p and q are approximately doubled each time from $p = 6, q = 1$ to $p = 50, q = 10$. One sees a performance gain of about 0.5 dB from $p = 6, q = 1$ to $p = 12, q = 2$ and a performance gain of about 0.4 dB from $p = 12, q = 2$ to $p = 25, q = 5$. The performance gain from $p = 25, q = 5$ to $p = 50, q = 10$ is very small, less than 0.2 dB for all the values of SNR considered. Therefore, $p = 25, q = 5$ is also a good design choice, particularly if one wants to minimize hardware. In realistic hardware designs, one may need to tradeoff p with q. For example, consider a case where one can process and store 30 samples, so $p + q = 30$. Fig. 2.4 shows the performances of $\hat{\rho}$-SDC at different combinations of values of p and q, where $p + q = 30$. One sees that $p = 22, q = 8$ and $p = 25, q = 5$ offer the best performances among all the combinations. This suggests that p can be traded with q if a large value of q is not desirable. However, neither p nor q should be too small when $p + q$ is fixed. If p (or q) is too small, q (or p) becomes unnecessarily large.

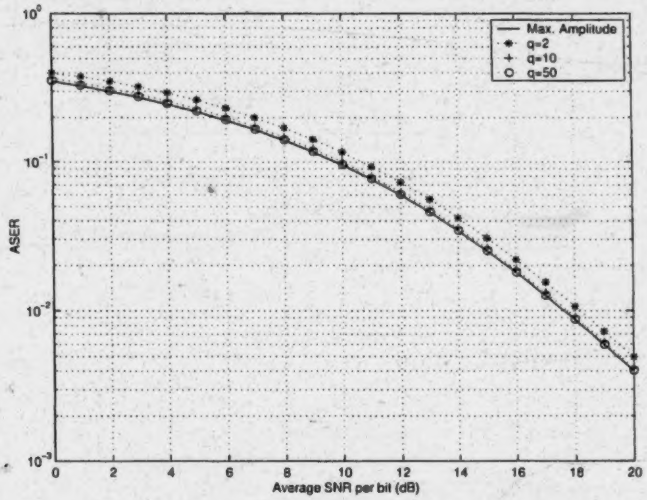

Figure 2.1. Comparison of the performances of $\hat{r}$-SDC (maximum estimated signal amplitude) and r-SDC (maximum signal amplitude) for different values of q when $N_2 = 5N_1$ and NC-BFSK is used.

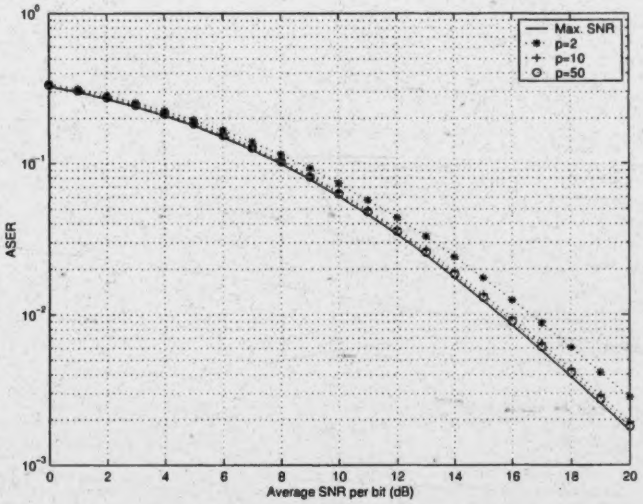

Figure 2.2. Comparison of the performances of $\hat{\rho}$-SDC (maximum estimated SNR) and ρ-SDC (maximum SNR) for different values of p when $q = 10$, $N_2 = 5N_1$ and NC-BFSK is used.

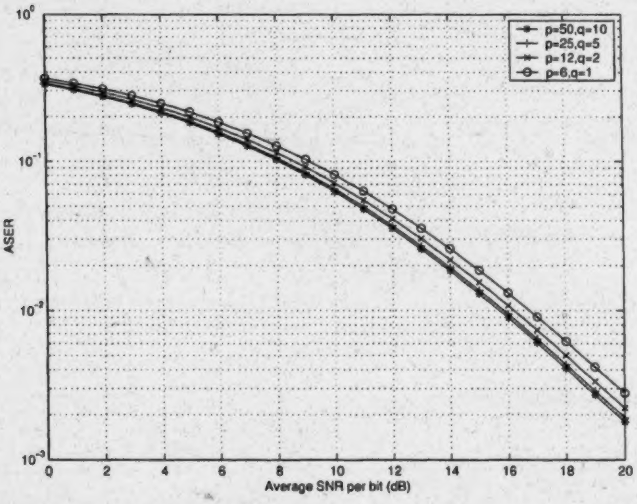

Figure 2.3. Comparison of the performances of $\hat{\rho}$-SDC (maximum estimated SNR) for selected values of p and q when $N_2 = 5N_1$ and NC-BFSK is used.

Furthermore, the performance loss due to the decrease of p (or q) cannot be fully compensated by that gained in the increase of q (or p) when p (or q) is too small.

Figs. 2.5 and 2.6 examine the effect of different noise power assumptions on the performance. To clarify this effect, both p and q are fixed at 2. Fig. 2.5 shows the performances of $\hat{r}$-SDC and $\hat{\rho}$-SDC when $N_2 = N_1$. As can be seen, when the noise powers on the two branches are the same, $\hat{r}$-SDC performs better than $\hat{\rho}$-SDC. The difference is about 0.9 dB in SNR when the ASER is 10^{-2}. This is expected, since $\hat{\rho}$-SDC introduces additional estimation errors in the estimation of the noise power, while $\hat{r}$-SDC avoids this by assuming balanced noise powers, which is the case when $N_2 = N_1$. Note also that r-SDC and ρ-SDC have the same performance, as the maximum SNR criterion is equivalent to the maximum signal amplitude criterion when the noise powers are equal and the estimation is perfect. Fig. 2.6 shows the performances of $\hat{r}$-SDC and $\hat{\rho}$-SDC when $N_2 = 5N_1$. From this figure, it can be seen that $\hat{\rho}$-SDC has a performance advantage of about 1.1 dB over $\hat{r}$-SDC, over the range of SNR values shown, when $N_2 = 5N_1$. This is expected, as in this

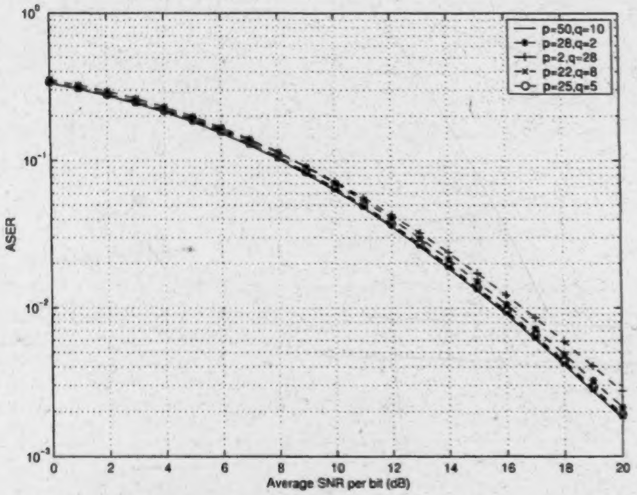

Figure 2.4. Comparison of the performances of $\hat{\rho}$-SDC (maximum estimated SNR) with different combinations of values of p and q when $p+q=30$, $N_2=5N_1$ and NC-BFSK is used.

case, the assumption of balanced noise powers used by $\hat{r}$-SDC is no longer true, and one has to take the unbalanced noise powers into account to minimize the probability of error. Comparing the performances of ρ-SDC with that of r-SDC, one sees that there is a penalty of about 2 dB when the ASER is 10^{-2}, caused by not accounting for the difference in the noise powers.

In conclusion, the preceding results have shown that $\hat{r}$-SDC performs better than $\hat{\rho}$-SDC when the noise powers are balanced, confirming intuitive reasoning. When the noise powers are unbalanced, $\hat{\rho}$-SDC performs better than $\hat{r}$-SDC. The performance penalty caused by the estimation errors decreases as p and/or q increase, and the performance penalty caused by ignoring noise power differences increases as the noise power difference increases. Using these results, the effects of channel state parameter estimation error and noise power imbalance can be evaluated quantitively.

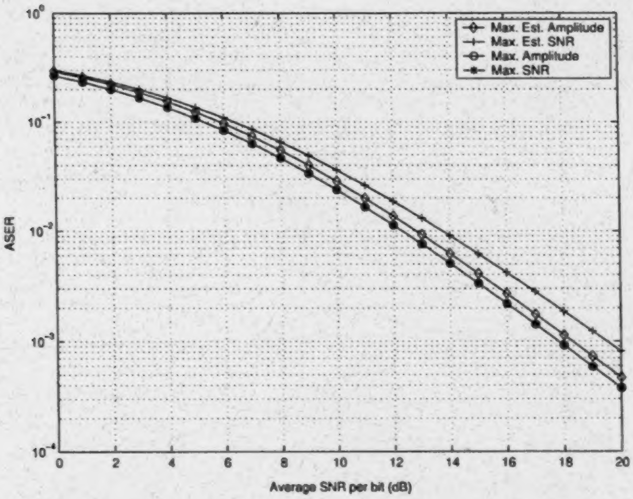

Figure 2.5. Performances of $\hat{r}$-SDC (maximum estimated signal amplitude), $\hat{\rho}$-SDC (maximum estimated SNR), r-SDC (maximum signal amplitude), and ρ-SDC (maximum SNR) at $p = 2$ and $q = 2$ with $N_2 = N_1$ for NC-BFSK signaling.

2.3 Receiver Designs Using Error Statistics of Channel State Parameter Estimation

In the previous section, we have evaluated the effect of channel estimation errors on the performances of SDCs. In this section, instead of analyzing the effect of estimation errors, we use estimation error statistics to design better receivers. The estimation error statistics are obtained from analysis or simulation of estimator performances, as in Section 2.1. Previous works include the following. In [4], the authors derived a diversity receiver that outperforms maximal ratio combiner (MRC) for binary equal energy signals on Ricean fading channels with Gaussian channel estimation errors. In [51], [52] and [53], the authors discussed optimal reception of MPSK, M-ary quadrature amplitude modulation (MQAM), and coded MQAM signals with additive Gaussian channel estimation errors on single Ricean fading channel, respectively. Here, we derive new diversity receivers for

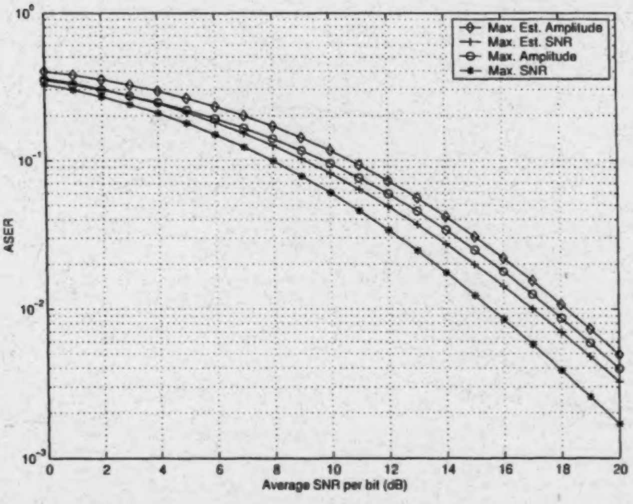

Figure 2.6. Performances of $\hat{r}$-SDC (maximum estimated signal amplitude), $\hat{\rho}$-SDC (maximum estimated SNR), r-SDC (maximum signal amplitude), and ρ-SDC (maximum SNR) at $p = 2$ and $q = 2$ with $N_2 = 5N_1$ for NC-BFSK signaling.

arbitrary M-ary signals on L_d independent and identically distributed channels, under more general assumptions on the knowledge of the channel state parameter estimates. Two important cases are studied. The first case applies to either Nakagami-m or Ricean fading channel, while the second case applies to Ricean fading channel only.

We consider L_d independent and identically distributed branches. Each is slowly and flatly fading. The received signal on the l-th branch is matched filtered and normalized with respect to the transmitted signal energy to give $y_l = A_l s_k + n_l$, where $A_l = A_l^R + jA_l^I$ is the complex channel gain on the l-th branch, s_k is the transmitted signaling point, and n_l is the AWGN on the l-th branch with mean zero and variance $2\sigma^2$. The conditional PDF of the received signal is

$$f(\mathbf{y}|\mathbf{A}^R, \mathbf{A}^I, s_k) = \frac{1}{(2\pi\sigma^2)^{L_d}} \cdot e^{-\frac{(\mathbf{y}-s_k\mathbf{A})(\mathbf{y}-s_k\mathbf{A})^H}{2\sigma^2}} \tag{2.108}$$

where $\mathbf{y} = (y_1 \quad y_2 \quad \cdots \quad y_{L_d})$, $\mathbf{A}^R = (A_1^R \quad A_2^R \quad \cdots \quad A_{L_d}^R)$, $\mathbf{A}^I = (A_1^I \quad A_2^I \quad \cdots \quad A_{L_d}^I)$, $\mathbf{A} = \mathbf{A}^R + j\mathbf{A}^I$, and $(\cdot)^H$ is the conjugate transpose. If A_l is perfectly known, one can derive the optimal diversity receiver from (2.108) as [54]

$$\hat{k} = \arg\min_{k} \left\{ (\mathbf{y} - s_k \mathbf{A})(\mathbf{y} - s_k \mathbf{A})^H \right\} \tag{2.109}$$

which implements MRC. However, the value of A_l is usually unavailable and has to be estimated. Denote $\hat{A}_l = \hat{A}_l^R + j\hat{A}_l^I$, $\hat{\mathbf{A}}^R = (\hat{A}_1^R \quad \hat{A}_2^R \quad \cdots \quad \hat{A}_{L_d}^R)$, $\hat{\mathbf{A}}^I = (\hat{A}_1^I \quad \hat{A}_2^I \quad \cdots \quad \hat{A}_{L_d}^I)$, and $\hat{\mathbf{A}} = \hat{\mathbf{A}}^R + j\hat{\mathbf{A}}^I$ as the channel estimates obtained from a channel estimator. The "intuitive" structure

$$\hat{k} = \arg\min_{k} \left\{ (\mathbf{y} - s_k \hat{\mathbf{A}})(\mathbf{y} - s_k \hat{\mathbf{A}})^H \right\} \tag{2.110}$$

is frequently used in practice.

2.3.1 Case 1

In the first case, one only has knowledge of the conditional PDF of $\hat{A}_l$, conditioned on A_l, as

$$f(\hat{A}_l | A_l) = \frac{1}{2\pi\tilde{\alpha}^2} e^{-\frac{1}{2\tilde{\alpha}^2}|\hat{A}_l - A_l - b|^2} \tag{2.111}$$

which is a Gaussian distribution with mean $A_l + b$ and variance $2\tilde{\alpha}^2$. This occurs, for example, in the case when the channel gain is estimated using the ML estimator in Section 2.1 and there may be some decision errors in the estimation, or a robust channel estimator in [51] and there may be some synchronization errors in the estimation. This also occurs in the case when a large sample size is used to estimate A_l and the estimate is asymptotically Gaussian distributed [15]. The values of b and $2\tilde{\alpha}^2$ can be determined through performance analysis or simulation of the estimator. Without any prior knowledge of $\mathbf{A}^R$ and $\mathbf{A}^I$, assume that they are uniformly distributed over $(-R, R)$, where R is a large positive real number. Simulation results show that this assumption leads to a detector with good performance. The joint PDF of the channel gain estimate and the true channel gain satisfies $f(\hat{\mathbf{A}}^R, \hat{\mathbf{A}}^I, \mathbf{A}^R, \mathbf{A}^I) = D \cdot f(\hat{\mathbf{A}}^R, \hat{\mathbf{A}}^I | \mathbf{A}^R, \mathbf{A}^I)$ where $f(\hat{\mathbf{A}}^R, \hat{\mathbf{A}}^I | \mathbf{A}^R, \mathbf{A}^I) = \prod_{l=1}^{L_d} f(\hat{A}_l | A_l)$, as $f(\mathbf{A}^R, \mathbf{A}^I) = D$ is a constant and $f(\hat{\mathbf{A}}^R, \hat{\mathbf{A}}^I, \mathbf{A}^R, \mathbf{A}^I) = f(\hat{\mathbf{A}}^R, \hat{\mathbf{A}}^I | \mathbf{A}^R, \mathbf{A}^I) f(\mathbf{A}^R, \mathbf{A}^I)$. Assume that $\hat{A}_l$ is independent of n_l (for example, $\hat{A}_l$ is estimated using pilot symbols). Using the average likelihood ratio test (ALRT) principle, the likelihood function can be derived as $f(\mathbf{y}, \hat{\mathbf{A}}^R, \hat{\mathbf{A}}^I | s_k) = \int f(\mathbf{y} | \mathbf{A}^R, \mathbf{A}^I, s_k) \cdot$

$f(\hat{\mathbf{A}}^R, \hat{\mathbf{A}}^I, \mathbf{A}^R, \mathbf{A}^I) d\mathbf{A}^R d\mathbf{A}^I$. Solving the integral, one has

$$f(\mathbf{y}, \hat{\mathbf{A}}^R, \hat{\mathbf{A}}^I | s_k) = \frac{D_1}{(\tilde{\alpha}^2 |s_k|^2 + \sigma^2)^{L_d}} e^{-\frac{[\mathbf{y} - s_k(\hat{\mathbf{A}} - b)][\mathbf{y} - s_k(\hat{\mathbf{A}} - b)]^H}{2(\tilde{\alpha}^2 |s_k|^2 + \sigma^2)}} \tag{2.112}$$

where D_1 is a constant independent of s_k. The new diversity receiver is derived from (2.112) as

$$\hat{k} = \arg\min_k \left\{ \phi_1^{(k)} + \frac{[\mathbf{y} - s_k(\hat{\mathbf{A}} - b)][\mathbf{y} - s_k(\hat{\mathbf{A}} - b)]^H}{2(\tilde{\alpha}^2 |s_k|^2 + \sigma^2)} \right\} \tag{2.113}$$

where $\phi_1^{(k)} = L_d \ln(\tilde{\alpha}^2 |s_k|^2 + \sigma^2)$. Several observations can be made from (2.113). First, the receiver in (2.113) is general and applies to all channel estimators giving a channel estimate based on the conditional PDF in (2.111). Second, the receiver in (2.113) can be used on either a Nakagami-m fading channel and a Ricean fading channel, as the distribution of A_l is not specified and used to derive (2.113). Finally, in general, the receiver in (2.113) is not equivalent to that in (2.110). However, when constant modulus signaling is used and the channel estimator is unbiased, (2.113) will be equivalent to (2.110). In this case, (2.113) won't provide any improvement over (2.110). By omitting the bias term $\phi_1^{(k)}$ in (2.113), a simplified structure of (2.113) is derived from (2.113) as

$$\hat{k} = \arg\min_k \left\{ \frac{[\mathbf{y} - s_k(\hat{\mathbf{A}} - b)][\mathbf{y} - s_k(\hat{\mathbf{A}} - b)]^H}{2(\tilde{\alpha}^2 |s_k|^2 + \sigma^2)} \right\}. \tag{2.114}$$

The receiver in (2.114) can also be derived by applying the generalized likelihood ratio test (GLRT) principle [17]. The GLRT receiver assumes that A_l is a deterministic but unknown number, while the ALRT receiver assumes that A_l is a random number [14]. As will be seen later, the GLRT receiver outperforms the ALRT receiver for small values of SNR, and it underperforms the ALRT receiver for large values of SNR.

A special case results when the ML estimator in Section 2.1 is used and no decision errors occur, or when the robust channel estimator [51] is used and no synchronization errors occur. In this case, $\hat{A}_l = A_l + g_l$ where g_l is a Gaussian channel estimation error with mean zero and variance $2\sigma_e^2$. One has $b = 0$ and $\tilde{\alpha}^2 = \sigma_e^2$ as $E\{\hat{A}_l | A_l\} = A_l$ and $Var\{\hat{A}_l | A_l\} = 2\sigma_e^2$. Thus, the receiver in (2.113) specializes to

$$\hat{k} = \arg\min_k \left\{ \psi_1^{(k)} + \frac{(\mathbf{y} - s_k\hat{\mathbf{A}})(\mathbf{y} - s_k\hat{\mathbf{A}})^H}{2(\lambda |s_k|^2 + 1)\sigma^2} \right\} \tag{2.115}$$

and the receiver in (2.114) specializes to

$$\hat{k} = \arg\min_{k} \left\{ \frac{(\mathbf{y}-s_k\hat{\mathbf{A}})(\mathbf{y}-s_k\hat{\mathbf{A}})^H}{2(\lambda|s_k|^2+1)} \right\} \tag{2.116}$$

where $\psi_1^{(k)} = L_d \ln(\lambda|s_k|^2+1)$ and $\lambda = \frac{\sigma_e^2}{\sigma^2}$ is the ratio of the channel estimation error variance to the noise variance, a constant independent of the SNR for fixed channel estimators [51], [52].

2.3.2 Case 2

In this case, in addition to knowing the conditional PDF of $\hat{A}_l$, one also knows the PDF of A_l, and therefore, the joint PDF of $\hat{A}_l$ and A_l. A diversity receiver that is better than the conventional MRC for binary equal energy signals when $\hat{A}_l$ and A_l are jointly Gaussian distributed has been proposed previously in [4]. We extend it to arbitray M-ary signals. Similar to [4, eq. (11-1-11)], their joint PDF can be written as

$$f(\hat{A}_l^R,\hat{A}_l^I,A_l^R,A_l^I) = \frac{1}{4\pi^2\Delta}\cdot e^{-\frac{1}{2}(\mathbf{A}-\mathbf{e})\mathbf{M}^{-1}(\mathbf{A}-\mathbf{e})^T} \tag{2.117}$$

where $\mathbf{A} = (\hat{A}_l^R \quad \hat{A}_l^I \quad A_l^R \quad A_l^I)$, $\mathbf{e} = E\{\mathbf{A}\} = (\hat{e}^R \quad \hat{e}^I \quad e^R \quad e^I)$, $\mathbf{M}$ is the covariance matrix with $\mathbf{M} = \begin{bmatrix} \tilde{\alpha}^2 & 0 & R_c & R_s \\ 0 & \tilde{\alpha}^2 & -R_s & R_c \\ R_c & -R_s & \alpha^2 & 0 \\ R_s & R_c & 0 & \alpha^2 \end{bmatrix}$, $\Delta = |\mathbf{M}|^{\frac{1}{2}}$ is the square root of the determinant of $\mathbf{M}$, $\tilde{\alpha}^2$ is the variance of $\hat{A}_l^R$ (or $\hat{A}_l^I$), α^2 is the variance of A_l^R (or A_l^I), R_c is the covariance of $\hat{A}_l^R$ and A_l^R (or $\hat{A}_l^I$ and A_l^I), R_s is the covariance of $\hat{A}_l^R$ and A_l^I, and $\hat{A}_l^R$ and A_l^R are assumed independent of $\hat{A}_l^I$ and A_l^I, respectively. Solving $f(\mathbf{y},\hat{\mathbf{A}}^R,\hat{\mathbf{A}}^I|s_k) = \int f(\mathbf{y}|\mathbf{A}^R,\mathbf{A}^I,s_k)\cdot f(\hat{\mathbf{A}}^R,\hat{\mathbf{A}}^I,\mathbf{A}^R,\mathbf{A}^I)d\mathbf{A}^R d\mathbf{A}^I$ where $f(\hat{\mathbf{A}}^R,\hat{\mathbf{A}}^I,\mathbf{A}^R,\mathbf{A}^I) = \prod_{l=1}^{L_d} f(\hat{A}_l^R,\hat{A}_l^I,A_l^R,A_l^I)$, one has

$$\begin{aligned} f(\mathbf{y},\hat{\mathbf{A}}^R,\hat{\mathbf{A}}^I|s_k) &= \frac{D_2}{\left[(\Delta|s_k|^2+\tilde{\alpha}^2\sigma^2)\right]^{L_d}}\cdot e^{-\frac{|s_k|^2[\alpha^2\hat{\mathbf{A}}\hat{\mathbf{A}}^H+(c_1\hat{\mathbf{A}}^R+c_2\hat{\mathbf{A}}^I)\mathbf{I}^H+c_0L_d]}{2(\Delta|s_k|^2+\tilde{\alpha}^2\sigma^2)}} \\ &\cdot e^{-\frac{\sigma^2(\hat{\mathbf{A}}-\hat{e})(\hat{\mathbf{A}}-\hat{e})^H+\tilde{\alpha}^2\mathbf{y}\mathbf{y}^H}{2(\Delta|s_k|^2+\tilde{\alpha}^2\sigma^2)}}\cdot e^{-\frac{Re\{\mathbf{y}s_k^*[(c_3+jc_4)\mathbf{I}-2(R_c-jR_s)\hat{\mathbf{A}}]^H\}}{2(\Delta|s_k|^2+\tilde{\alpha}^2\sigma^2)}} \end{aligned} \tag{2.118}$$

where D_2 is a constant independent of s_k, $\mathbf{I}$ is a $1\times L$ all-one row vector, $\hat{e} = \hat{e}^R + j\hat{e}^I$, $c_0 = (\hat{e}^R)^2\alpha^2 + (\hat{e}^I)^2\alpha^2 + (e^R)^2\tilde{\alpha}^2 + (e^I)^2\tilde{\alpha}^2 - 2R_c\hat{e}^Re^R - 2R_c\hat{e}^Ie^I - 2R_s\hat{e}^Re^I + 2R_s\hat{e}^Ie^R$, $c_1 = -2\hat{e}^R\alpha^2 + 2R_ce^R +$

$2R_s e^I$, $c_2 = -2\hat{e}^I \alpha^2 + 2R_c e^I - 2R_s e^R$, $c_3 = -2e^R \tilde{\alpha}^2 + 2R_c \hat{e}^R - 2R_s \hat{e}^I$, and $c_4 = -2e^I \tilde{\alpha}^2 + 2R_c \hat{e}^I + 2R_s \hat{e}^R$. A new diversity receiver in this case is derived from (2.118) as

$$\begin{aligned} \hat{k} &= \arg\min_k \left\{ \frac{|s_k|^2[\alpha^2 \hat{\mathbf{A}}\hat{\mathbf{A}}^H + (c_1 \hat{\mathbf{A}}^R + c_2 \hat{\mathbf{A}}^I + c_0)\mathbf{I}^H]}{2(\Delta |s_k|^2 + \tilde{\alpha}^2 \sigma^2)} + \frac{\sigma^2(\hat{\mathbf{A}} - \hat{e})(\hat{\mathbf{A}} - \hat{e})^H + \tilde{\alpha}^2 \mathbf{y}\mathbf{y}^H}{2(\Delta |s_k|^2 + \tilde{\alpha}^2 \sigma^2)} \right. \\ &+ \left. \frac{Re\{\mathbf{y} s_k^* [(c_3 + jc_4)\mathbf{I} - (R_c - jR_s)\hat{\mathbf{A}}]^H\}}{2(\Delta |s_k|^2 + \tilde{\alpha}^2 \sigma^2)} + \phi_2^{(k)} \right\} \end{aligned} \tag{2.119}$$

where $\phi_2^{(k)} = L_d \ln(\Delta |s_k|^2 + \tilde{\alpha}^2 \sigma^2)$. The diversity receiver in (2.119) is a maximum likelihood structure and is optimum for the case that the joint density of the channel gains and the channel gain estimates are given by (2.117). The structure of this receiver is general. It includes the case when the channel gain estimate is corrupted by temporal decorrelation [45] plus estimator noise as well as the case when the estimation error is independent of the true channel gain. It can, thus, model the case when the channel gain estimate is obtained from a pilot carrier [45], pilot symbol [55], or a robust channel estimator [51]. Note that the receiver in (2.119) requires more knowledge of the channel estimate statistics than those in (2.113) and (2.114). As a result, it generally has a more limited usage than those in (2.113) and (2.114). For example, when the channel is Ricean faded and the Gaussian distribution of the true channel gain is known, all three diversity receivers in (2.113), (2.114) and (2.119) can be used. However, when the channel is Ricean faded and the Gaussian distribution of the true channel gain is unknown, or when the channel is Nakagami-*m* faded and the true channel gain follows an unknown non-Gaussian distribution, only the diversity receivers in (2.113) and (2.114) can be used. Therefore, Case 1 is not a special case of Case 2. The former assumes less knowledge of the channel statistics than the latter and, consequently, accommodates a wider usage. Finally, when the ML estimator in Section 2.1 and the robust channel estimator [51] are used, one has $\tilde{\alpha}^2 = \alpha^2 + \sigma_e^2$, $R_c = \alpha^2$, $R_s = 0$, $\Delta = \alpha^2 \sigma_e^2$, $\hat{e}^R = e^R$ and $\hat{e}^I = e^I$. Thus,

$$\begin{aligned} \hat{k} &= \arg\min_k \left\{ \psi_2^{(k)} + \frac{(\hat{\mathbf{A}} - \hat{e})(\hat{\mathbf{A}} - \hat{e})^H}{2(\frac{\alpha^2}{\sigma^2}\lambda |s_k|^2 + \frac{\alpha^2}{\sigma^2} + \lambda)\sigma^2} \right. \\ &+ \left. \frac{\frac{\alpha^2}{\sigma^2}(\mathbf{y} - s_k \hat{\mathbf{A}})(\mathbf{y} - s_k \hat{\mathbf{A}})^H + \lambda(\mathbf{y} - \hat{e} s_k)(\mathbf{y} - \hat{e} s_k)^H}{2(\frac{\alpha^2}{\sigma^2}\lambda |s_k|^2 + \frac{\alpha^2}{\sigma^2} + \lambda)\sigma^2} \right\} \end{aligned} \tag{2.120}$$

where $\psi_2^{(k)} = L_d \ln(\frac{\alpha^2}{\sigma^2}\lambda |s_k|^2 + \frac{\alpha^2}{\sigma^2} + \lambda)$.

2.3.3 Comparison of Case 1 and Case 2

Fig. 2.7 shows the symbol error rates (SERs) of perfect MRC in (2.109), MRC with estimation errors in (2.110) and the new diversity receiver in (2.120) for binary phase shift keying (BPSK) signaling with $L_d = 2$. The true channel gain is assumed to follow a known Gaussian distribution with mean $e = 2 + j2$ and variance 2 so that all the new diversity receivers derived in the previous section can be fairly compared. Assume that the ML estimator in Section 2.1 is used. Note that the receivers in (2.115) and (2.116) are equivalent to that in (2.110) in this case. One sees that the gain of the new diversity receiver in (2.120) over MRC with estimation errors in (2.110) is negligible when $\lambda = 0.1$. When $\lambda = 0.4$, at SER = 10^{-2}, the new diversity receiver has a performance gain of about 0.6 dB over MRC with estimation errors. The gain decreases as the SNR increases. Fig. 2.8 shows the performances of perfect MRC in (2.109), MRC with estimation errors in (2.110) and the new diversity receiver in (2.120) for BPSK signaling with $L_d = 4$. Again, the performance gain of the new diversity receiver in (2.120) over MRC with estimation errors in (2.110) is negligible when $\lambda = 0.1$. When $\lambda = 0.4$, the performance gain is about 1.2 dB at SER = 10^{-2}.

Fig. 2.9 compares the performances of the diversity receivers in (2.109), (2.110), (2.115) (2.116), and (2.120) for 16-ary quadrature amplitude modulation (16-QAM) signaling with $L_d = 2$. When the SER is 10^{-2}, the gain of the new diversity receiver in (2.120) over MRC with estimation errors in (2.110) is about 0.2 dB for $\lambda = 0.1$ and about 1.0 dB for $\lambda = 0.4$. Thus, the improvement for 16-QAM signaling is greater than that for BPSK signaling. Comparing the new diversity receiver in (2.115) to MRC with estimation errors at a SER of 10^{-2}, the gain is negligible when $\lambda = 0.1$ and about 0.4 dB when $\lambda = 0.4$. Therefore, Case 1 leads to less improvement than Case 2. This is expected, as Case 1 assumes less knowledge of the channel estimate statistics than Case 2. Finally, comparing the receiver in (2.116) to MRC with estimation errors, one observes that the receiver in (2.116) is better than MRC with estimation errors at small values of SNR, and is worse than MRC with estimation errors at large values of SNR. Since (2.116) is a suboptimal structure, it is not necessarily better than (2.110). Fig. 2.10 compares the performances of the diversity receivers

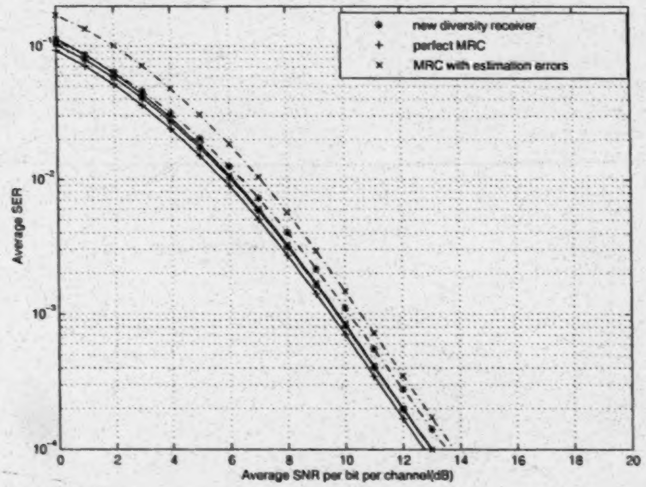

Figure 2.7. Performances of perfect MRC in (2.109), MRC with estimation errors in (2.110) and the new diversity receiver in (2.120) for BPSK signaling at $\lambda = 0.1$ (solid line) and $\lambda = 0.4$ (dash-dotted line) when $L_d = 2$ in Ricean fading channels.

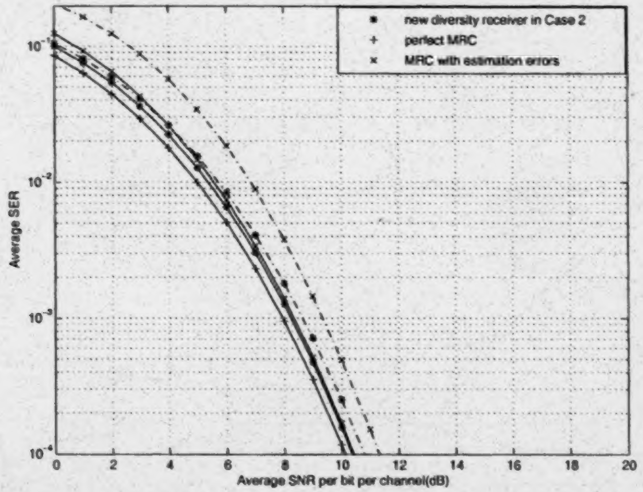

Figure 2.8. Performances of perfect MRC in (2.109), MRC with estimation errors in (2.110) and the new diversity receiver in (2.120) for BPSK signaling at $\lambda = 0.1$ (solid line) and $\lambda = 0.4$ (dash-dotted line) when $L_d = 4$ in Ricean fading channels.

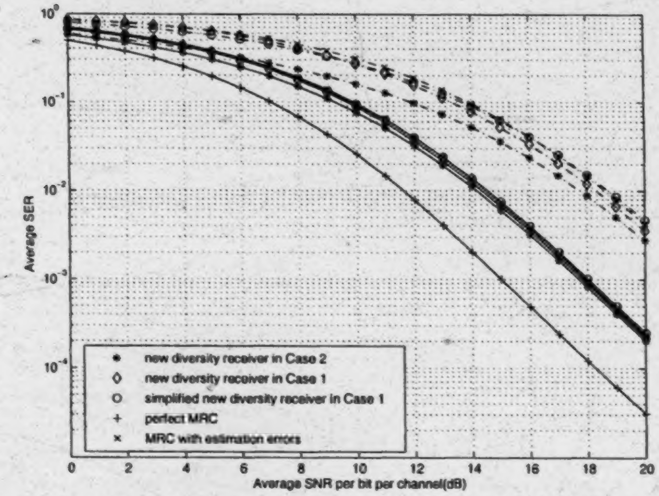

Figure 2.9. Performances of perfect MRC in (2.109), MRC with estimation errors in (2.110), the new diversity receivers in (2.115), (2.116) and (2.120) for 16-QAM signaling at $\lambda = 0.1$ (solid line) and $\lambda = 0.4$ (dash-dotted line) when $L_d = 2$ in Ricean fading channels.

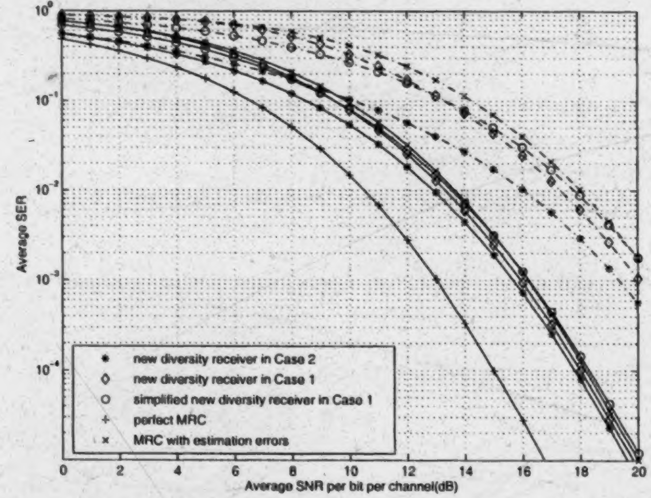

Figure 2.10. Performances of perfect MRC in (2.109), MRC with estimation errors in (2.110), the new diversity receivers in (2.115), (2.116) and (2.120) for 16-QAM signaling at $\lambda = 0.1$ (solid line) and $\lambda = 0.4$ (dash-dotted line) when $L_d = 4$ in Ricean fading channels.

in (2.109), (2.110), (2.115) (2.116), and (2.120) for 16-QAM signaling with $L_d = 4$. When the SER is 10^{-2}, the gain of the new diversity receiver in (2.120) over MRC with estimation errors is about 0.7 dB for $\lambda = 0.1$ and about 2.0 dB for $\lambda = 0.4$. Also, comparing the new diversity receiver in (2.115) to MRC with estimation errors at a SER of 10^{-2}, the gain is about 0.2 dB when $\lambda = 0.1$ and about 0.6 dB when $\lambda = 0.4$. Similar observations to those from Fig. 2.9 can be made from Fig. 2.10. Note that the performance gains of the new diversity receivers are achieved at the cost of more complex structures. Also, exact values of the channel estimate statistics are assumed. Note further that the performances of MRC with estimation errors in Figs. 2.7 to 2.10 don't show error rate floors, since we only consider slow fading channels here and the estimation error variance of the channel estimator used in the simulation varies inversely with the SNR.

2.4 Receiver Designs Using Pilot Symbols of Channel State Parameter Estimation

In the previous section, we have derived new receivers by using statistics of the channel estimation errors. In practical communication systems, channel estimation is often performed with the aid of known pilot symbols. The statistics of the channel estimation errors are usually functions of the known pilot symbols. In this work, instead of using statistics of the channel estimation errors determined by the pilot symbols, we design novel receivers that use the pilot symbols directly. Specifically, we design optimum receivers for pilot symbol assisted modulation (PSAM) signals in Rayleigh and Ricean fading channels [55]- [57].

2.4.1 System Model

We assume that the autocovariance function of the fading process is known, as was assumed in [55] and [57]. Consider a PSAM system where symbols are transmitted in frames of length Q. Without loss of generality, assume that, in each frame, the first symbol is a pilot symbol and the following $Q-1$ symbols are data symbols. Each data symbol comes from a set of M possible signals, $\{b_j\}_{j=1}^{M}$.

The pilot symbol comes from the same signaling set, but its value is known as $\tilde{b}$. These frames are transmitted over a flatly fading channel. The received signal can be written as

$$y(t) = A(t)s(t) + n(t) \tag{2.121}$$

where $s(t)$ is the transmitted signal, $A(t)$ is the complex channel gain, and $n(t)$ is the AWGN. Following the model in [55], the transmitted signal, $s(t)$, satisfies

$$s(t) = \sum_{l=-\infty}^{\infty} b_l p(t - lT) \tag{2.122}$$

where b_l is the value of the l-th symbol coming from $\{b_j\}_{j=1}^{M}$, T is the symbol period, and $p(t)$ is the shaping pulse with energy E_p. If the l-th symbol is a pilot symbol, $b_l = \tilde{b}$ is known. Otherwise, it is unknown and may be one of M possible values. The complex channel gain, $A(t)$, is a Gaussian random process. Denote it as $A(t) = A^R(t) + jA^I(t)$. If the channel is Rayleigh faded, one has [54]

$$E\{A^R(t)\} = E\{A^I(t)\} = 0 \tag{2.123a}$$

$$Cov\left(A^R(t), A^I(t+\tau)\right) = 0 \tag{2.123b}$$

$$Cov\left(A^R(t), A^R(t+\tau)\right) = Cov\left(A^I(t), A^I(t+\tau)\right) = \alpha^2 \tilde{R}(\tau) \tag{2.123c}$$

where $Cov\left(A^R(t), A^R(t+\tau)\right) = E\{[A^R(t) - E\{A^R(t)\}][A^R(t+\tau) - E\{A^R(t+\tau)\}]\}$ is the autocovariance of $A^R(t)$, $Cov\left(A^I(t), A^I(t+\tau)\right) = E\{[A^I(t) - E\{A^I(t)\}][A^I(t+\tau) - E\{A^I(t+\tau)\}]\}$ is the autocovariance of $A^I(t)$; $Cov\left(A^R(t), A^I(t+\tau)\right) = E\{[A^R(t) - E\{A^R(t)\}][A^I(t+\tau) - E\{A^I(t+\tau)\}]\}$ is the cross-covariance between $A^R(t)$ and $A^I(t)$, and $\tilde{R}(\tau)$ is the normalized autocovariance function with $\tilde{R}(0) = 1$. If the channel is Ricean faded, one then has [54]

$$E\{A^R(t)\} = e^R(t) \tag{2.124a}$$

$$E\{A^I(t)\} = e^I(t) \tag{2.124b}$$

$$Cov\left(A^R(t), A^I(t+\tau)\right) = 0 \tag{2.124c}$$

$$Cov\left(A^R(t), A^R(t+\tau)\right) = Cov\left(A^I(t), A^I(t+\tau)\right) = \alpha^2 \tilde{R}(\tau). \tag{2.124d}$$

Note that (2.123) is a special case of (2.124) when $e^R(t) = 0$ and $e^I(t) = 0$. If the scattering in the Rayleigh or Ricean channel is isotropic, one further has [54]

$$\tilde{R}(\tau) = J_0(2\pi f_D \tau) \tag{2.125}$$

where f_D is the maximum Doppler shift in the channel. Although our analysis is not limited to any specific $\tilde{R}(\tau)$, (2.125) will be used to obtain illustrative examples. The noise $n(t)$ is also a Gaussian random process. It has mean zero and autocovariance $E\{n(t)n^*(t+\tau)\} = N_0\delta(\tau)$.

Similar to [55], it is assumed that no intersymbol interference occurs. The received signal in (2.121) is matched filtered and sampled at the time $t = lT$. The received signal sample of the l-th symbol is

$$y_l = A(lT)b_l E_p + n_l \tag{2.126}$$

where $A(lT)$ is the Gaussian channel gain sample with mean zero (on a Rayleigh fading channel) or $e^R(lT) + je^I(lT)$ (on a Ricean fading channel) and variance $2\alpha^2$, and n_l is a Gaussian noise sample with mean zero and variance $2\sigma^2 = N_0 E_p$. The average SNR per bit is derived from (2.126) as

$$\bar{\rho}_b = \frac{E_p^2 \Omega}{2\sigma^2 \log_2 M} \cdot \frac{|\tilde{b}|^2 + (Q-1)E\{|b_l|^2\}}{Q-1}. \tag{2.127}$$

where $\Omega = E\{A(lT)A^*(lT)\}$ is the mean power of the fading and $E\{|b_l|^2\}$ is the average transmitted signal energy. This measure accounts for the power penalty caused by sending pilot symbols.

Without loss of generality, let the 0-th symbol in the transmitted sequence be a pilot symbol. Then, the 1-st, 2-nd, $\cdots$, $(Q-1)$-th symbols in the transmitted sequence are data symbols in the same frame. Assume that the J nearest pilot symbols are used to assist the data symbol detection in a frame. Let the function, $\lfloor x \rfloor$, return the largest integer that is less than or equal to x. In this case, the $(\lfloor -\frac{J-1}{2} \rfloor Q)$-th, $\cdots$, 0-th, $\cdots$, $(\lfloor \frac{J-1}{2} \rfloor Q)$-th symbols in the transmitted sequence are the pilot symbols that will be used to assist the detection of the 1-st, 2-nd, $\cdots$, $(Q-1)$-th symbols in the transmitted sequence, which are data symbols. For clarity, it is desirable to use different notations to denote the received signal for the data symbol and the received signals for the pilot symbols. Denote the received signal sample of the data symbol to be detected as

$$y_k = A_k b_k + n_k \tag{2.128}$$

where $A_k = A(kT)E_p$ and $k = 1,2,\cdots,Q-1$. Also, denote the received signal samples of the pilot symbols that will be used to assist the detection of b_k as

$$p_i = B_i\tilde{b} + n_i \tag{2.129}$$

where $B_i = A(iT)E_p$ and $i = \lfloor -\frac{J-1}{2} \rfloor Q, \cdots, \lfloor \frac{J-1}{2} \rfloor Q$. Express the complex channel gains as $A_k = A_k^R + jA_k^I$ and $B_i = B_i^R + jB_i^I$. It can be derived from (2.123) and (2.124) that the covariance between A_k^R and B_i^R (or A_k^I and B_i^I) is $w_k(i) = E_p^2\alpha^2\tilde{R}((k-i)T)$, $i = \lfloor -\frac{J-1}{2} \rfloor Q, \cdots, \lfloor \frac{J-1}{2} \rfloor Q$, and the covariance between A_i^R and B_j^R (or A_i^I and B_j^I) is $C_k(i,j) = E_p^2\alpha^2\tilde{R}((i-j)T)$, $i,j = \lfloor -\frac{J-1}{2} \rfloor Q, \cdots, \lfloor \frac{J-1}{2} \rfloor Q$. Finally, introduce the notations $\mathbf{w}_k = [w_k(\lfloor -\frac{J-1}{2} \rfloor Q) \quad \cdots \quad w_k(\lfloor \frac{J-1}{2} \rfloor Q)]$ and $\mathbf{C}_k = \{C_k(i,j)\}$, where $C_k(i,j)$ is the (i,j)-th element of $\mathbf{C}_k$, for later use.

By using the assumptions and notations defined above, it can be derived from (2.128) that the conditional PDF of y_k, conditioned on A_k and b_k, is

$$f(y_k|A_k,b_k) = \frac{1}{2\pi\sigma^2}\exp\left\{-\frac{1}{2\sigma^2}|y_k - A_k b_k|^2\right\}. \tag{2.130}$$

Similarly, the conditional PDF of p_i, conditioned on B_i, can be derived from (2.129) as

$$f(p_i|B_i) = \frac{1}{2\pi\sigma^2}\exp\left\{-\frac{1}{2\sigma^2}|p_i - B_i\tilde{b}|^2\right\}. \tag{2.131}$$

Since the symbol-spaced noise samples are independent, from (2.130) and (2.131), the conditional joint PDF of y_k and p_i ($i = \lfloor -\frac{J-1}{2} \rfloor Q, \cdots, \lfloor \frac{J-1}{2} \rfloor Q$), conditioned on A_k, b_k and B_i ($i = \lfloor -\frac{J-1}{2} \rfloor Q, \cdots, \lfloor \frac{J-1}{2} \rfloor Q$), can be obtained as

$$f(y_k,\mathbf{p}|A_k,\mathbf{B},b_k) = \frac{1}{(2\pi\sigma^2)^{J+1}}\exp\left\{-\frac{1}{2\sigma^2}|y_k - A_k b_k|^2 - \frac{1}{2\sigma^2}\sum_{i=\lfloor -\frac{J-1}{2} \rfloor Q}^{\lfloor \frac{J-1}{2} \rfloor Q}|p_i - B_i\tilde{b}|^2\right\} \tag{2.132}$$

where $\mathbf{p} = [p_{\lfloor -\frac{J-1}{2} \rfloor Q} \quad \cdots \quad p_{\lfloor \frac{J-1}{2} \rfloor Q}]$ and $\mathbf{B} = [B_{\lfloor -\frac{J-1}{2} \rfloor Q} \quad \cdots \quad B_{\lfloor \frac{J-1}{2} \rfloor Q}]$. Finally, the likelihood function for joint processing of the data symbol and the pilot symbols can be derived by solving

$$f(y_k,\mathbf{p}|b_k) = \int\int f(y_k,\mathbf{p}|A_k,\mathbf{B},b_k)\cdot f(A_k,\mathbf{B})dA_k d\mathbf{B} \tag{2.133}$$

where $f(A_k,\mathbf{B})$ is the joint PDF of A_k and $\mathbf{B}$. The optimum PSAM (OPSAM) signal detector is obtained by maximizing (2.133) with respect to b_k.

2.4.2 Optimum PSAM Signal Detectors

To derive the value of b_k that maximizes (2.133), one needs to solve the integration in (2.133) first. Since the joint PDF of A_k and $\mathbf{B}$, $f(A_k, \mathbf{B})$, depends on the fading channel model, we examine the optimum maximum likelihood detector for the PSAM signal on Rayleigh and Ricean fading channels separately in the sequel.

2.4.2.1 Rayleigh Fading Channel

In a Rayleigh fading channel, the statistics of the fading process are determined by (2.123). Thus, the joint PDF of A_k and $\mathbf{B}$ can be derived as

$$f(A_k, \mathbf{B}) = \frac{1}{(2\pi)^{J+1}|\mathbf{H}_k|} e^{-\frac{1}{2}\mathbf{A}_R \mathbf{H}_k^{-1} \mathbf{A}_R^T - \frac{1}{2}\mathbf{A}_I \mathbf{H}_k^{-1} \mathbf{A}_I^T} \tag{2.134}$$

where T denotes the transpose of a matrix or a vector, $|\mathbf{H}_k|$ denotes the determinant of $\mathbf{H}_k$, $\mathbf{H}_k^{-1}$ denotes the inverse of $\mathbf{H}_k$, $\mathbf{A}_R = [A_k^R \quad B^R_{\lfloor -\frac{J-1}{2} \rfloor Q} \quad \cdots \quad B^R_{\lfloor \frac{J-1}{2} \rfloor Q}]$ is a $1 \times (J+1)$ row vector consisting of the real components of the channel gains, $\mathbf{A}_I = [A_k^I \quad B^I_{\lfloor -\frac{J-1}{2} \rfloor Q} \quad \cdots \quad B^I_{\lfloor \frac{J-1}{2} \rfloor Q}]$ is a $1 \times (J+1)$ row vector consisting of the imaginary components of the channel gains, $\mathbf{H}_k$ is the $(J+1) \times (J+1)$ covariance matrix with

$$\mathbf{H}_k = \begin{bmatrix} E_p^2 \alpha^2 & \mathbf{w}_k \\ \mathbf{w}_k^T & \mathbf{C}_k \end{bmatrix}, \tag{2.135}$$

and $\mathbf{w}_k$, $\mathbf{C}_k$ are defined as before. By using (2.134) in (2.133) and solving the resulting integral, it is shown in Appendix C that

$$f(y_k, \mathbf{p}|b_k) = \frac{1}{(2\pi\sigma^2)^{J+1}|\mathbf{H}_k||\mathbf{F}_k|} e^{\frac{1}{2}\mathbf{u}\mathbf{F}_k^{-1}\mathbf{u}^T + \frac{1}{2}\mathbf{v}\mathbf{F}_k^{-1}\mathbf{v}^T - \frac{|y_k|^2}{2\sigma^2} - \frac{\sum_{i=\lfloor -\frac{J-1}{2} \rfloor Q}^{\lfloor \frac{J-1}{2} \rfloor Q} |p_i|^2}{2\sigma^2}} \tag{2.136}$$

where $\mathbf{F}_k = \mathbf{H}_k^{-1} + \mathbf{G}_k$, $\mathbf{G}_k = \frac{1}{\sigma^2}\begin{bmatrix} |b_k|^2 & \mathbf{0} \\ \mathbf{0}^T & |\tilde{b}|^2\mathbf{E} \end{bmatrix}$, $\mathbf{u} = [\frac{Re\{y_k b_k^*\}}{\sigma^2} \quad \frac{Re\{p_{\lfloor -\frac{J-1}{2} \rfloor Q}\tilde{b}^*\}}{\sigma^2} \quad \cdots \quad \frac{Re\{p_{\lfloor \frac{J-1}{2} \rfloor Q}\tilde{b}^*\}}{\sigma^2}]$, $\mathbf{v} = [\frac{Im\{y_k b_k^*\}}{\sigma^2} \quad \frac{Im\{p_{\lfloor -\frac{J-1}{2} \rfloor Q}\tilde{b}^*\}}{\sigma^2} \quad \cdots \quad \frac{Im\{p_{\lfloor \frac{J-1}{2} \rfloor Q}\tilde{b}^*\}}{\sigma^2}]$, $\mathbf{0}$ is a $1 \times J$ zero vector, and $\mathbf{E}$ is a $J \times J$ identity matrix. The optimum maximum likelihood detector chooses the value of b_k that maximizes (2.136) from a set of M signals, $\{b_j\}_{j=1}^M$. Two important special cases will be discussed next.

If the energies of the M possible transmitted signals are equal, such as those in MPSK signaling, $|b_k|^2$ is a constant independent of k. This implies that $|b_k|^2 = |\bar{b}|^2$ and that the energy of the signal does not affect the choice of b_k in (2.136). Ignoring those terms independent of b_k, one has

$$f(y_k, \mathbf{p}|b_k) \propto e^{\frac{1}{2}\mathbf{u}\mathbf{F}_k^{-1}\mathbf{u}^T + \frac{1}{2}\mathbf{v}\mathbf{F}_k^{-1}\mathbf{v}^T}. \tag{2.137}$$

Further simplification shows that

$$f(y_k, \mathbf{p}|b_k) \propto e^{\frac{1}{\sigma^4}\left\{Re\{y_k b_k^*\}Re\{\bar{b}^*\mathbf{p}\cdot\mathbf{S}_k^T\} + Im\{y_k b_k^*\}Im\{\bar{b}^*\mathbf{p}\cdot\mathbf{S}_k^T\}\right\}} \tag{2.138}$$

where $\mathbf{S}_k$ is a $1 \times J$ vector derived in Appendix D as

$$\mathbf{S}_k = \mathbf{w}_k \cdot \left[\left(\frac{\sigma^2}{\sigma^2 + E_p^2\alpha^2|b_k|^2} + \frac{\sigma^2|b_k|^2\mathbf{w}_k\mathbf{Z}_1\mathbf{w}_k^T}{(\sigma^2 + E_p^2\alpha^2|b_k|^2)^2}\right)\mathbf{E} - \frac{\sigma^2}{\sigma^2 + E_p^2\alpha^2|b_k|^2}\mathbf{Z}_1 \cdot \mathbf{C}_k\right] \tag{2.139}$$

and

$$\mathbf{Z}_1 = \left[\mathbf{C}_k + \frac{\sigma^2}{|\bar{b}|^2}\mathbf{E} - \frac{|b_k|^2}{\sigma^2 + E_p^2\alpha^2|b_k|^2}\mathbf{w}_k^T\mathbf{w}_k\right]^{-1}. \tag{2.140}$$

Finally, the OPSAM signal detector in the Rayleigh fading channel when the transmitted signals are of equal energies can be obtained from (2.138) as

$$\hat{b}_k = arg \max_{b_k \in \{b_j\}_{j=1}^M} \{Re\{y_k b_k^* X_k^*\}\} \tag{2.141}$$

where $X_k = \bar{b}^*\mathbf{p} \cdot \mathbf{S}_k^T$ and $\mathbf{S}_k$ is given by (2.139). Note that the optimum detector in (2.141) is actually a correlator which weights the received symbol signal, r_k, with the conjugate channel gain estimate, X_k^*, and then correlates the compensated received symbol signal with the corresponding signal value, b_k, to make the data decision.

The performance of the OPSAM signal detector for BPSK signaling in the Rayleigh fading channel can be analyzed as follows. It has been derived in [1] that the bit error rate (BER) of any BPSK signal detector satisfying [1, eq. (B-1)] is [1, eq. (B-21)]

$$P_b = \frac{v_1}{v_1 + v_2}, \tag{2.142}$$

where $v_1 = \frac{1}{\sqrt{c_{yy}c_{XX}} + c_{yX}}$, $v_2 = \frac{1}{\sqrt{c_{yy}c_{XX}} - c_{yX}}$, c_{yy} is the variance of y_k, c_{XX} is the variance of X_k, and c_{yX} is the covariance between y_k and X_k. Denoting $c_1 = \frac{c_{yX}}{\sqrt{c_{yy}c_{XX}}}$ as the covariance coefficient between y_k

and X_k, (2.142) can be rewritten as

$$P_b = \frac{1}{2}(1 - c_1). \tag{2.143}$$

If the OPSAM signal detector is used, it can be shown that $c_{yy} = 2(E_p^2\alpha^2 + \sigma^2)$, $c_{XX} = 2|\tilde{b}|^2\mathbf{S}_k \cdot (|\tilde{b}|^2\mathbf{C}_k + \sigma^2\mathbf{E}) \cdot \mathbf{S}_k^T$, and $c_{yX} = 2|\tilde{b}|^2\mathbf{w}_k \cdot \mathbf{S}_k^T$. Then,

$$c_1 = \frac{|\tilde{b}|^2\mathbf{w}_k \cdot \mathbf{S}_k^T}{\sqrt{(E_p^2\alpha^2 + \sigma^2)|\tilde{b}|^2\mathbf{S}_k \cdot (|\tilde{b}|^2\mathbf{C}_k + \sigma^2\mathbf{E}) \cdot \mathbf{S}_k^T}} \tag{2.144}$$

where $\mathbf{S}_k$ is given by (2.139). Therefore, the BER of the BPSK signaling can be evaulated analytically using (2.143) and (2.144).

If the energies of the M possible transmitted signals are not equal, such as those in MQAM signaling, the choice of b_k in (2.136) depends on $|b_k|^2$ as well. In this case, one has

$$f(y_k, \mathbf{p}|b_k) \propto \frac{1}{|\mathbf{F}_k|} e^{\frac{1}{2}\mathbf{u}\mathbf{F}_k^{-1}\mathbf{u}^T + \frac{1}{2}\mathbf{v}\mathbf{F}_k^{-1}\mathbf{v}^T}, \tag{2.145}$$

as $|\mathbf{F}_k|$ is related to $|b_k|^2$. Examination of $\mathbf{F}_k$ shows that

$$|\mathbf{F}_k| = \left|\frac{|b_k|^2}{\sigma^2} + d\right| \cdot \left|\mathbf{Z}_2 + \frac{|\tilde{b}|^2}{\sigma^2}\mathbf{E}\right| \tag{2.146}$$

where $d = \frac{1}{E_p^2\alpha^2} + \frac{1}{E_p^4\alpha^4}\mathbf{w}_k \cdot \mathbf{Z}_2 \cdot (\frac{\sigma^2}{|\tilde{b}|^2}\mathbf{Z}_2 + \mathbf{E})^{-1}\mathbf{w}_k^T$ and $\mathbf{Z}_2 = [\mathbf{C}_k - \frac{1}{E_p^2\alpha^2}\mathbf{w}_k^T\mathbf{w}_k]^{-1}$. Substituting (2.146) in (2.145) and doing some additional simplifications, one has

$$f(y_k, \mathbf{p}|b_k) \propto \frac{1}{|\frac{|b_k|^2}{\sigma^2} + d|} e^{\frac{1}{\sigma^4}\left\{Re\{y_k b_k^*\}Re\{\tilde{b}^*\mathbf{p}\cdot\mathbf{S}_k^T\} + Im\{y_k b_k^*\}Im\{\tilde{b}^*\mathbf{p}\cdot\mathbf{S}_k^T\}\right\} + W_k} \tag{2.147}$$

where $W_k = \frac{|r_k|^2|b_k|^2 f_k}{2\sigma^4} + \frac{|\tilde{b}|^2 Re\{\mathbf{p}\mathbf{Z}_3\mathbf{p}^H\}}{2\sigma^4}$, $(\cdot)^H$ denotes the conjugate transpose, $\mathbf{Z}_3 = \mathbf{C}_k - \frac{|b_k|^2\mathbf{w}_k^T\mathbf{w}_k}{\sigma^2 + E_p^2\alpha^2|b_k|^2}(1 + \frac{|b_k|^2\mathbf{w}_k\mathbf{Z}_1\mathbf{w}_k^T}{\sigma^2 + E_p^2\alpha^2|b_k|^2}) + \frac{|b_k|^2(\mathbf{C}_k\mathbf{Z}_1\mathbf{w}_k^T\mathbf{w}_k + \mathbf{w}_k^T\mathbf{w}_k\mathbf{Z}_1\mathbf{C}_k)}{\sigma^2 + E_p^2\alpha^2|b_k|^2} - \mathbf{C}_k\mathbf{Z}_1\mathbf{C}_k$, and $f_k = \frac{E_p^2\alpha^2\sigma^2}{\sigma^2 + E_p^2\alpha^2|b_k|^2} - \frac{\sigma^4\mathbf{w}_k\mathbf{Z}_1\mathbf{w}_k^T}{(\sigma^2 + E_p^2\alpha^2|b_k|^2)^2}$. Finally, the OPSAM signal detector in the Rayleigh fading channel when the transmitted signals have unequal energies can be derived from (2.147) as

$$\hat{b_k} = arg \max_{b_k \in \{b_j\}_{j=1}^M} \left\{Re\{y_k b_k^* X_k^*\} - \sigma^4 \ln\left|\frac{|b_k|^2}{\sigma^2} + d\right| + \sigma^4 W_k\right\} \tag{2.148}$$

where $X_k = \tilde{b}^*\mathbf{p}\cdot\mathbf{S}_k^T$ and $\mathbf{S}_k$ is given by (2.139), as before. Comparing (2.148) with (2.141), one sees that (2.148) has two additional bias terms caused by the unequal energies of the transmitted signals, as expected.

2.4.2.2 Ricean Fading Channel

In the previous subsection, we derived the OPSAM signal detector for the Rayleigh fading channel. This exposition served to develop the theory. In practice, some real-world channels exhibit Ricean fading. In this subsection, we will derive the OPSAM signal detector for the Ricean fading channel.

In the Ricean fading channel, the statistics of the fading process are determined by (2.124). Therefore, the joint PDF of A_k and $\mathbf{B}$ satisfies

$$f(A_k, \mathbf{B}) = \frac{1}{(2\pi)^{J+1}|\mathbf{H}_k|} e^{-\frac{1}{2}(\mathbf{A}_R-\mathbf{e}_R)\mathbf{H}_k^{-1}(\mathbf{A}_R-\mathbf{e}_R)^T-\frac{1}{2}(\mathbf{A}_I-\mathbf{e}_I)\mathbf{H}_k^{-1}(\mathbf{A}_I-\mathbf{e}_I)^T} \tag{2.149}$$

where $\mathbf{e}_R = E\{\mathbf{A}_R\} = [e_k^R \quad e^R_{\lfloor -\frac{J-1}{2} \rfloor Q} \quad \cdots \quad e^R_{\lfloor \frac{J-1}{2} \rfloor Q}]$ and $\mathbf{e}_I = E\{\mathbf{A}_I\} = [e_k^I \quad e^I_{\lfloor -\frac{J-1}{2} \rfloor Q} \quad \cdots \quad e^I_{\lfloor \frac{J-1}{2} \rfloor Q}]$. One sees that (2.134) is a special case of (2.149) when $\mathbf{e}_R = \mathbf{0}$ and $\mathbf{e}_I = \mathbf{0}$. Similarly, by using (2.149) in (2.133) and solving the resulting integral, as shown in Appendix C, one can derive

$$f(y_k, \mathbf{p}|b_k) = \frac{1}{(2\pi\sigma^2)^{J+1}|\mathbf{H}_k||\mathbf{F}_k|} e^{\frac{1}{2}\mathbf{u}'\mathbf{F}_k^{-1}\mathbf{u}'^T+\frac{1}{2}\mathbf{v}'\mathbf{F}_k^{-1}\mathbf{v}'^T-\frac{1}{2}\mathbf{e}_R\mathbf{H}_k^{-1}\mathbf{e}_R^T-\frac{1}{2}\mathbf{e}_I\mathbf{H}_k^{-1}\mathbf{e}_I^T-\frac{|y_k|^2}{2\sigma^2}-\frac{\sum_{i=\lfloor -\frac{J-1}{2}\rfloor Q}^{\lfloor -\frac{J-1}{2}\rfloor Q}|p_i|^2}{2\sigma^2}} \tag{2.150}$$

where $\mathbf{u}' = \mathbf{u} + \mathbf{e}_R\mathbf{H}_k^{-1}$ and $\mathbf{v}' = \mathbf{v} + \mathbf{e}_I\mathbf{H}_k^{-1}$. The OPSAM signal detector in the Ricean fading channel is derived by maximizing (2.150) with respect to b_k. As previously, two important special cases are discussed.

Again, we begin with the case when the transmitted signals have equal energies. In this case, $\mathbf{F}_k$ is independent of b_k. Thus,

$$f(y_k, \mathbf{p}|b_k) \propto e^{\frac{1}{2}\mathbf{u}'\mathbf{F}_k^{-1}\mathbf{u}'^T+\frac{1}{2}\mathbf{v}'\mathbf{F}_k^{-1}\mathbf{v}'^T}. \tag{2.151}$$

By using $\mathbf{u}' = \mathbf{u} + \mathbf{e}_R\mathbf{H}_k^{-1}$ and $\mathbf{v}' = \mathbf{v} + \mathbf{e}_I\mathbf{H}_k^{-1}$ in (2.151), one has

$$f(y_k, \mathbf{p}|b_k) \propto e^{\frac{1}{\sigma^4}\{Re\{y_k b_k^*\}Re\{\tilde{b}^*\mathbf{p}\cdot\mathbf{S}_k^T\}+Im\{y_k b_k^*\}Im\{\tilde{b}^*\mathbf{p}\cdot\mathbf{S}_k^T\}+\sigma^2 Re\{y_k b_k^*\}\mathbf{e}_R\cdot\mathbf{Q}_k^T+\sigma^2 Im\{y_k b_k^*\}\mathbf{e}_I\cdot\mathbf{Q}_k^T\}} \tag{2.152}$$

where $\mathbf{Q}_k$ is a $1 \times (J+1)$ vector derived in Appendix D as

$$\mathbf{Q}_k = \left[\frac{\sigma^2}{\sigma^2 + E_p^2\alpha^2|b_k|^2}\left(1 + \frac{|b_k|^2\mathbf{w}_k\mathbf{Z}_1\mathbf{w}_k^T}{\sigma^2 + E_p^2\alpha^2|b_k|^2}\right) \quad -\frac{\sigma^2}{\sigma^2 + E_p^2\alpha^2|b_k|^2}\mathbf{w}_k\mathbf{Z}_1\right]. \tag{2.153}$$

The OPSAM signal detector for equal energy signals in the Ricean fading channel is derived from (2.152) as

$$\hat{b_k} = arg \max_{b_k \in \{b_j\}_{j=1}^{M}} \{Re\{y_k b_k^* V_k^*\}\} \tag{2.154}$$

where $V_k = X_k + \sigma^2 Y_k$, $Y_k = (\mathbf{e}_R + j\mathbf{e}_I)\mathbf{Q}_k^T$, and $\mathbf{Q}_k$ is given by (2.153). Several observations can be made from (2.154). First, note that Y_k in V_k is a deterministic number. Thus, V_k and X_k have different means but the same variances. Second, by comparing (2.154) with (2.141), one sees that there is an additional term in (2.154) caused by the non-zero specular component in the Ricean fading channel. If the channel is Rayleigh faded, $\mathbf{e}_R = \mathbf{e}_I = \mathbf{0}$ and (2.154) will specialize to (2.141).

The performance of the OPSAM signal detector for BPSK signaling in the Ricean fading channel can also be derived by using results in [1]. The BER is [1, eq. (B-21)]

$$P_b = Q_1(a,b) - \frac{v_2/v_1}{1+v_2/v_1} I_0(ab) e^{-\frac{1}{2}(a^2+b^2)}, \tag{2.155}$$

where $Q_1(\cdot,\cdot)$ is the Marcum's Q function, $I_n(\cdot)$ is the n-th order modified Bessel function of the first kind, $a = \frac{1}{\sqrt{2}} |\frac{\bar{y}_k}{\sqrt{c_{yy}}} - \frac{\bar{V}_k}{\sqrt{c_{VV}}}|$, $b = \frac{1}{\sqrt{2}} |\frac{\bar{y}_k}{\sqrt{c_{yy}}} + \frac{\bar{V}_k}{\sqrt{c_{VV}}}|$, $\bar{y}_k = E_p(e_k^R + je_k^I)$, $\bar{V}_k = |\tilde{b}|^2 E_p(\mathbf{e}_p^R + j\mathbf{e}_p^I)\mathbf{S}_k^T + \sigma^2 Y_k$, $\mathbf{e}_p^R = [e^R_{\lfloor -\frac{L-1}{2} \rfloor Q} \quad \cdots \quad e^R_{\lfloor -\frac{L-1}{2} \rfloor Q}]$, $\mathbf{e}_p^I = [e^I_{\lfloor -\frac{L-1}{2} \rfloor Q} \quad \cdots \quad e^I_{\lfloor \frac{L-1}{2} \rfloor Q}]$, $c_{VV} = c_{XX}$, $c_{yV} = c_{yX}$, and v_1, v_2, c_{yy}, c_{XX}, c_{yX} are defined as before. Denoting $c_2 = \frac{c_{yV}}{\sqrt{c_{yy}c_{VV}}}$ as the covariance coefficient between y_k and V_k, (2.155) can be rewritten as

$$P_b = Q_1(a,b) - \frac{1+c_2}{2} I_0(ab) e^{-\frac{1}{2}(a^2+b^2)}. \tag{2.156}$$

It can be verified that $c_2 = c_1$ and (2.143) is a special case of (2.156) when the specular component in the fading channel is zero.

If the energies of the transmitted signals are not equal, the likelihood function in (2.150) can only be simplified to be

$$f(y_k, \mathbf{p}|b_k) \propto \frac{1}{|\mathbf{F}_k|} e^{\frac{1}{2}\mathbf{u}'\mathbf{F}_k^{-1}\mathbf{u}'^T + \frac{1}{2}\mathbf{v}'\mathbf{F}_k^{-1}\mathbf{v}'^T}. \tag{2.157}$$

Again, by using (2.146) and the expressions of $\mathbf{u}'$ and $\mathbf{v}'$ in (2.157), one has

$$f(y_k, \mathbf{p}|b_k) \propto \frac{1}{|\frac{|b_k|^2}{\sigma^2} + d|} e^{\frac{1}{\sigma^2}\{Re\{y_k b_k^* V_k^*\}\} + W_k + T_k} \tag{2.158}$$

where $T_k = \frac{1}{2}\mathbf{e}_R\mathbf{H}_k^{-1}\mathbf{F}_k^{-1}\mathbf{H}_k^{-1}\mathbf{e}_R^T + \frac{1}{2}\mathbf{e}_I\mathbf{H}_k^{-1}\mathbf{F}_k^{-1}\mathbf{H}_k^{-1}\mathbf{e}_I^T + \frac{Re\{\tilde{b}(\mathbf{e}_R+j\mathbf{e}_I)\mathbf{D}_k^T\mathbf{p}^H\}}{\sigma^2}$, $\mathbf{D}_k = [-\frac{|b_k|^2\mathbf{Z}_4^T\mathbf{w}_k^T}{\sigma^2+E_p^2\alpha^2|b_k|^2} \quad \mathbf{Z}_4^T]$ and $\mathbf{Z}_4 = \frac{\sigma^2}{|\tilde{b}|^2}\mathbf{Z}_1$. Finally, the OPSAM signal detector for unequal energy signals in the Ricean fading channel is

$$\hat{b_k} = arg \max_{b_k \in \{b_j\}_{j=1}^M} \left\{ Re\{y_k b_k^* V_k^*\} - \sigma^4 \ln\left|\frac{|b_k|^2}{\sigma^2} + d\right| + \sigma^4 W_k + \sigma^4 T_k \right\}. \tag{2.159}$$

Comparing (2.159) with (2.154), one observes that there are three additional bias terms caused by the unequal energies of the transmitted signals in (2.159). Also, comparing (2.159) with (2.148), one sees that there is an additional bias term caused by the non-zero specular component in the Ricean fading channel. When the specular component is zero, the Ricean fading channel will specialize to the Rayleigh fading channel and (2.159) will specialize to (2.148), as expected.

2.4.3 Conventional PSAM Signal Detectors

In [55] and [57], the conventional PSAM (CPSAM) signal detector for BPSK signaling was derived. This detector obtains the channel gain estimate, X_k^I, by using a Wiener filter. Following the ideas in [55], [57] and using notations and symbols defined here, one can show that

$$X_k^I = \tilde{b}^* \mathbf{p} \cdot \mathbf{S'}_k^T \tag{2.160}$$

with

$$\mathbf{S'}_k = \mathbf{w}_k \cdot \left[|\tilde{b}|^2 \mathbf{C}_k + \sigma^2 \mathbf{E}\right]^{-1} \tag{2.161}$$

in a Rayleigh fading channel for BPSK signaling. Note that the Wiener filter given by (2.160) can also be used in a Rician fading channel [57]. This is the case when the specular component in the Ricean fading channel is unknown. This is also the case when the Ricean fading channel is non-stationary with a time-varying specular component and an optimum Wiener filter may be unavailable, as studied in this paper. Our OPSAM signal detector improves the channel gain estimate in (2.160) by using a joint processing of the data symbol and the pilot symbols in a Rayleigh fading channel and by using knowledge of the time-varying specular component and a joint processing of the data symbol and the pilot symbols in a Ricean fading channel, as can be seen from (2.141) and (2.154). Note further that one could also improve (2.160) by simply using knowledge of the time-varying specular component in the Wiener filter. This detector hasn't been derived previously in the literature, and it cannot be considered as the conventional detector for comparison. Moreover, the Wiener filter in a non-stationary Ricean channel with time-varying specular component may not be optimal. Our OPSAM signal detector is an optimal maximum likelihood structure in a non-stationary Rician channel that necessarily uses knowledge of the time-varying specular components and jointly processes the data symbol and the pilot symbols. Thus, there is no need to derive the Wiener filter using the time-varying specular component either. The channel gain estimate, $X_k^{\prime}$, is used to weight the received symbol signal for data decision. Therefore, the CPSAM signal detector for BPSK signaling can be written as [55]

$$\hat{b}_k = arg \max_{b_k \in \{b_j\}_{j=1}^{M}} \left\{ Re\{ y_k b_k^* X_k^{\prime *} \} \right\} \tag{2.162}$$

where $X_k^{\prime}$ is given by (2.160). Comparing (2.162) with (2.141) and (2.154), one sees that the CPSAM detector has similar computation complexity, as well as similar structure, to the OPSAM signal detectors. Both need matrix inversion to derive the channel estimate. The main difference comes from their ways of obtaining the channel gain estimate. In the CPSAM signal detector, the channel estimate is obtained by using the pilot symbols only, and channel estimation and signal detection are performed separately. However, in our OPSAM signal detectors, the channel estimate is obtained by processing both the pilot symbols and the data symbols, and channel estimation and signal detection are performed jointly. As a result, $X_k^{\prime}$ does not depend on b_k, while X_k and V_k do, in general.

The performance of the CPSAM signal detector for BPSK signaling can also be evaluated analytically by using results in [58]. The BER of the CPSAM signal detector for BPSK signaling in the Rayleigh fading channel is, again, given by (2.143). However, the value of c_1 in (2.144) should be replaced by [55]

$$c_1^{\prime} = \frac{|\tilde{b}|^2 \mathbf{w}_k \cdot \mathbf{S'}_k^T}{\sqrt{(E_p^2 \alpha^2 + \sigma^2)|\tilde{b}|^2 \mathbf{S'}_k \cdot (|\tilde{b}|^2 \mathbf{C}_k + \sigma^2 \mathbf{E}) \cdot \mathbf{S'}_k^T}} \tag{2.163}$$

for the CPSAM signal detector, where $\mathbf{S'}_k$ is defined in (2.161). It is proved in Appendix E that (2.144) and (2.163) are actually equivalent. Therefore, although the CPSAM signal detector and the OPSAM signal detector have different channel gain estimates, their BER performances are the same for BPSK signaling in the Rayleigh fading channel. We have confirmed numerically that the CPSAM channel gain estimate is a scaling of the OPSAM channel gain estimate for all cases considered. This implies that the CPSAM signal detector for BPSK signaling in Rayleigh fading channels is optimum in the sense of minimum probability of error. This fact has not previously been established.

The BER of the CPSAM signal detector for BPSK signaling in the Ricean fading channel can also be derived from (2.156). However, the values of a, b and c_2 in (2.156) should be replaced by the corresponding values of $a' = \frac{1}{\sqrt{2}}|\frac{\bar{y}_k}{\sqrt{c_{yy}}} - \frac{\bar{X}'_k}{\sqrt{c_{X'X'}}}|$, $b' = \frac{1}{\sqrt{2}}|\frac{\bar{y}_k}{\sqrt{c_{yy}}} + \frac{\bar{X}'_k}{\sqrt{c_{X'X'}}}|$, $c'_2 = c'_1$, where $\bar{X}'_k = |\tilde{b}|^2 E_p(\mathbf{e}_p^R + j\mathbf{e}_p^I)\mathbf{S}'^T_k$, $c_{X'X'} = |\tilde{b}|^2 \mathbf{S}'_k \cdot (|\tilde{b}|^2 \mathbf{C}_k + \sigma_n^2 \mathbf{E}) \cdot \mathbf{S}'^T_k$, and c'_1 is given in (2.163).

The CPSAM signal detector for 16-QAM signaling was derived in [55] and [56] as a threshold-based detector. To facilitate the performance comparison, we need its correlator-based form. Following similar ideas and procedures to those in [55] and [56], one can derive the CPSAM signal detector for 16-QAM signaling in its correlator-based form as

$$\hat{b}_k = arg \max_{b_k \in \{b_j\}_{j=1}^M} \left\{ Re\{y_k b_k^* X_k'^*\} - \frac{|b_k|^2}{2} |X'_k|^2 \right\} \tag{2.164}$$

for a Rayleigh fading channel, where X'_k is given in (2.160). Similarly, one can also use (2.164) in a Ricean fading channel. The CPSAM signal detector in this case is simpler than the OPSAM signal detector, as it has less bias terms and the channel gain estimate in (2.164) doesn't depend on the values of the possible transmitted signals.

2.4.4 Comparison of Optimum and Conventional PSAM Signal Detectors

Here, we compare the performances of the OPSAM signal detectors derived previously with those of the CPSAM signal detectors. For simplicity, we discuss the case when the means of the fading process in (2.124) are constant. The case of time-varying means can be examined accordingly. Thus, one has $e^R(t) = e^R$ and $e^I(t) = e^I$. Define $P^2 = (e^R)^2 + (e^I)^2$ as the local mean power of the line-of-sight component in the Ricean fading channel and $K = \frac{P^2}{2\alpha^2}$ as the Ricean K factor [1].

We examine the performances of the detectors at $K = 0$ (the Rayleigh fading channel), $K = 4$ and $K = 8$. Also, we assume that the scattering in the fading channel is isotropic, and the values of the normalized Doppler shift (normalized with respect to the symbol rate), $f_D T = 0.03$, $f_D T = 0.06$ and $f_D T = 0.09$ are used. The frame length is chosen to be $Q = 5$, and the number of pilot symbols used to assist the detection of the data symbol is chosen to be $J = 11$. The error rates are obtained by averaging the error rates of the data symbols over all positions in one frame.

Figs. 2.11 to 2.13 show the performances of the OPSAM signal detectors for BPSK signaling. One sees that the performance of the OPSAM signal detector improves when the power of the specular component in the fading channel increases or the normalized Doppler shift in the fading channel decreases. For example, when the $BER = 10^{-2}$ and $f_D T = 0.03$, the performance of the OPSAM signal detector for $K = 8$ is about 8.8 dB better than that for $K = 0$, and about 2.2 dB better than that for $K = 4$. When the $BER = 10^{-2}$ and $K = 0$, the performance of the OPSAM signal detector for $f_D T = 0.03$ is about 0.6 dB better than that for $f_D T = 0.06$, and about 2.6 dB better than that for $f_D T = 0.09$. This is expected, as a larger power in the specular component gives a better fading channel condition, and a smaller value of the normalized Doppler shift allows a more accurate channel gain estimate, which results in fewer errors in the data decisions.

Comparing the performance of the OPSAM signal detector with that of the CPSAM signal detector, one sees that the OPSAM signal detector has a performance gain over the conventional detector. The performance gain decreases when the power of the specular component in the channel decreases or the normalized Doppler shift in the channel decreases. As an example, when the $BER = 10^{-2}$ and $f_D T = 0.03$, the OPSAM signal detector has a performance gain of about 1.5 dB for $K = 8$, a performance gain of about 1.0 dB for $K = 4$, and no performance gain for $K = 0$. When the $BER = 10^{-2}$ and $K = 8$, the OPSAM signal detector has a performance gain of about 3.2 dB for $f_D T = 0.09$, about 2.4 dB for $f_D T = 0.06$, and about 1.5 dB for $f_D T = 0.03$. Observe that the performance gain decreases as $\bar{\rho}_b$ increases. This is explained as follows. Comparing (2.154) with (2.162), one sees that the performance gain of our OPSAM signal detector for BPSK signaling comes from the fact that we are using an additional offset, $\sigma^2 Y_k$, to calculate the channel

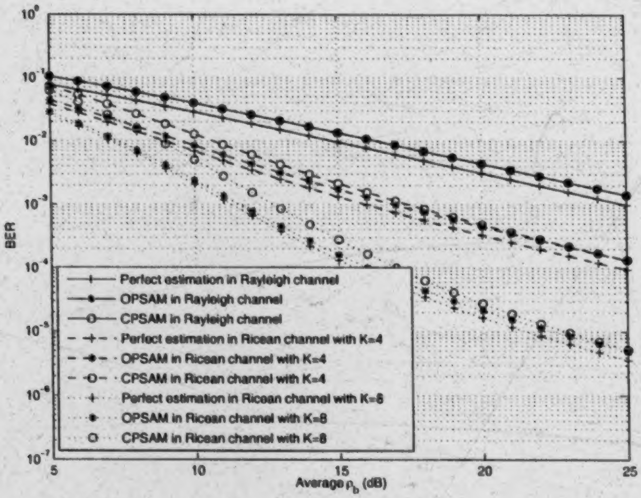

Figure 2.11. Performance comparison of the OPSAM signal detector and the CPSAM signal detector for BPSK signaling in Rayleigh and Ricean fading channels when $f_D T = 0.03$.

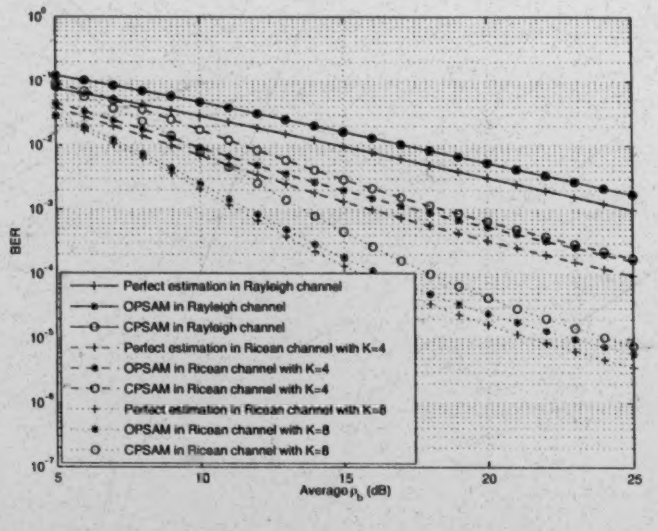

Figure 2.12. Performance comparison of the OPSAM signal detector and the CPSAM signal detector for BPSK signaling in Rayleigh and Ricean fading channels when $f_D T = 0.06$.

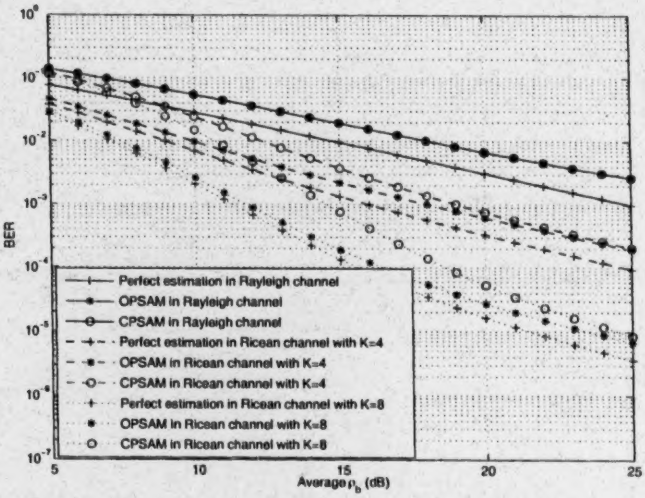

Figure 2.13. Performance comparison of the OPSAM signal detector and the CPSAM signal detector for BPSK signaling in Rayleigh and Ricean fading channels when $f_D T = 0.09$.

gain estimate. When the power of the specular component in the channel or the normalized Doppler shift in the channel decrease, or the SNR increases, Y_k or σ^2 will become relatively smaller, and the offset will become less significant. Then, the performance gain decreases.

Figs. 2.14 to 2.16 show the performances of the OPSAM signal detectors for 16-QAM signaling. Again, the performance of the OPSAM signal detector improves when the power of the specular component in the fading channel increases or the normalized Doppler shift in the fading channel decreases. At $SER = 10^{-2}$ and $f_D T = 0.03$, the OPSAM signal detector for $K = 8$ is about 10.0 dB better than that for $K = 0$, and about 2.6 dB better than that for $K = 4$. At $SER = 10^{-2}$ and $K = 0$, the OPSAM signal detector for $f_D T = 0.03$ is about 0.4 dB better than that for $f_D T = 0.06$, and about 0.8 dB better than that for $f_D T = 0.09$.

Comparing the performance of the OPSAM signal detector with that of the conventional detector, one sees that the optimum detector outperforms the conventional detector. When the $BER = 10^{-1}$ and $f_D T = 0.06$, the OPSAM signal detector has performance gains of about 1.0 dB for $K = 8$,

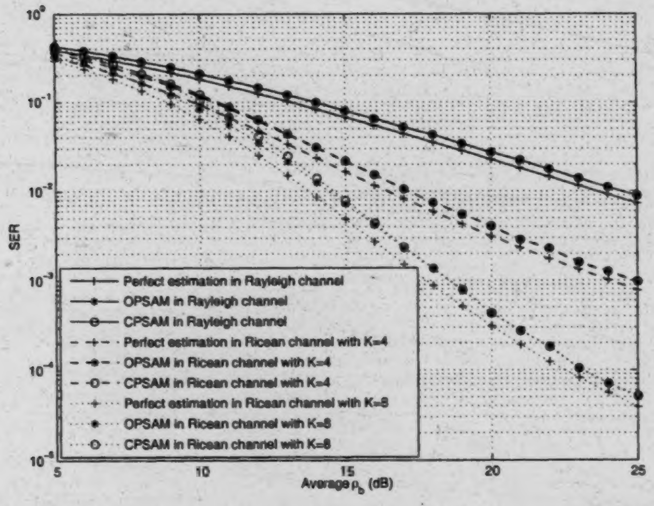

Figure 2.14. Performance comparison of the OPSAM signal detector and the CPSAM signal detector for 16-QAM signaling in Rayleigh and Ricean fading channels when $f_D T = 0.03$.

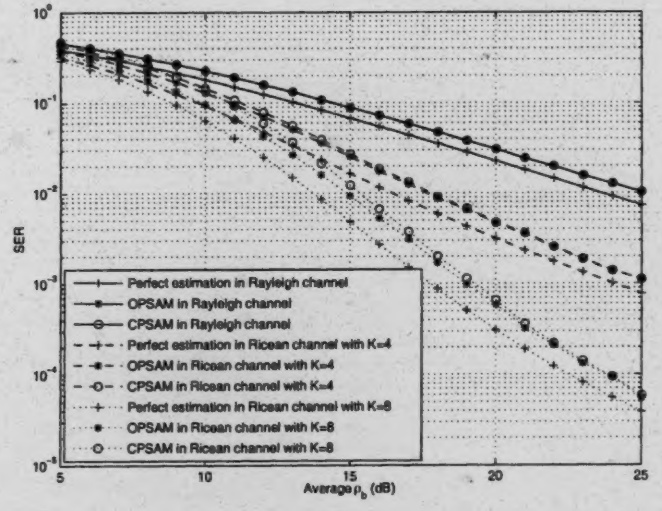

Figure 2.15. Performance comparison of the OPSAM signal detector and the CPSAM signal detector for 16-QAM signaling in Rayleigh and Ricean fading channels when $f_D T = 0.06$.

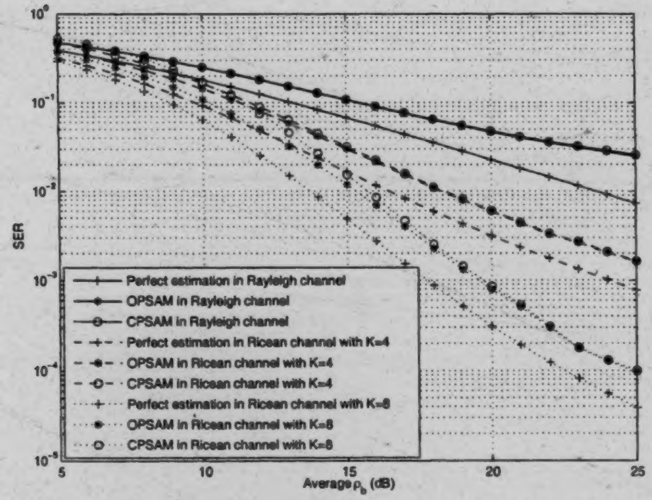

Figure 2.16. Performance comparison of the OPSAM signal detector and the CPSAM signal detector for 16-QAM signaling in Rayleigh and Ricean fading channels when $f_D T = 0.09$.

Table 2.1. Performance gains of OPSAM over CPSAM for BPSK

Fading Conditions $(f_D T, K)$	$SER = 10^{-1}$ (dB)	$SER = 10^{-2}$ (dB)	$SER = 10^{-3}$ (dB)	$SER = 10^{-4}$ (dB)
$f_D T = 0.03$, $K = 0$	0.0	0.0	0.0	0.0
$f_D T = 0.03$, $K = 4$	2.1	1.0	0.4	0.0
$f_D T = 0.03$, $K = 8$	2.5	1.5	1.2	0.9
$f_D T = 0.06$, $K = 0$	0.0	0.0	0.0	0.0
$f_D T = 0.06$, $K = 4$	2.7	1.6	0.8	0.0
$f_D T = 0.06$, $K = 8$	3.2	2.4	1.9	1.7
$f_D T = 0.09$, $K = 0$	0.0	0.0	0.0	0.0
$f_D T = 0.09$, $K = 4$	3.7	2.3	0.8	0.0
$f_D T = 0.09$, $K = 8$	4.3	3.2	2.9	2.1

Table 2.2. Performance gains of OPSAM over CPSAM for 16-QAM

Fading Conditions (f_DT,K)	$SER = 10^{-1}$ (dB)	$SER = 10^{-2}$ (dB)	$SER = 10^{-3}$ (dB)	$SER = 10^{-4}$ (dB)
$f_DT = 0.03, K = 0$	0.0	0.0	0.0	0.0
$f_DT = 0.03, K = 4$	0.1	0.0	0.0	0.0
$f_DT = 0.03, K = 8$	0.6	0.2	0.0	0.0
$f_DT = 0.06, K = 0$	0.0	0.0	0.0	0.0
$f_DT = 0.06, K = 4$	0.2	0.0	0.0	0.0
$f_DT = 0.06, K = 8$	1.0	0.3	0.1	0.0
$f_DT = 0.09, K = 0$	0.0	0.0	0.0	0.0
$f_DT = 0.09, K = 4$	0.4	0.0	0.0	0.0
$f_DT = 0.09, K = 8$	1.2	0.4	0.2	0.0

about 0.2 dB for $K = 4$, and approximately 0 dB for $K = 0$. When the $BER = 10^{-1}$ and $K = 8$, the OPSAM signal detector has a performance gain of about 1.2 dB for $f_DT = 0.09$, about 1.0 dB for $f_DT = 0.06$, and about 0.6 dB for $f_DT = 0.03$. Again, the performance gain decreases as the SNR increases. Note that the performance gains of the OPSAM signal detector over the conventional detector for 16-QAM signaling are smaller than the corresponding performance gains for BPSK signaling. Note further that our OPSAM signal detector requires knowledge of $\mathbf{e}_R$ and $\mathbf{e}_I$, or equivalently, P^2, as well as knowledge of $2\alpha^2$, as can be seen from (2.154) and (2.159), while the CPSAM signal detector only needs knowledge of $2\alpha^2$. Both $2\alpha^2$ and P^2 can be accurately estimated using estimators developed in [59]. Tables 2.1 and 2.2 show the performance gains in SNR of OPSAM over CPSAM for BPSK and 16-QAM, respectively. For BPSK signaling the gain ranges from 0 dB for Rayleigh fading to 4.3 dB for Rician fading with $f_DT = 0.09$ and $K = 8$ at an error rate of 10^{-1}. Observe that OPSAM has the desirable property that its gains over CPSAM are greatest at larger error rates, where the gains are most needed. The gain for BPSK signaling comes exclusively from

the use of P^2, while the large part of the gain for 16-QAM signaling comes from the use of P^2 and the joint processing of data and pilot symbols contributes some small gain. Next, we investigate the problem of channel state parameter estimation in a UWB system.

2.5 Channel State Parameter Estimation in a UWB System

Ultra-wide bandwidth technology has attracted much research interest recently [60]- [62]. In [63] and [64], ML estimators for the multipath gain and the multipath delay in a UWB system that use an isolated monocycle and an information-bearing signal, respectively, were proposed. In [65], the authors derived the CRLB for the DA ML channel estimator developed in [64]. In this work, we first derive closed-form expressions of the CRLBs for non-data-aided (NDA) ML channel gain estimation and channel delay estimation in a UWB system. Then, we propose some new NDA ML channel gain and delay estimators.

2.5.1 UWB System Model

We consider a time-hopping, pulse position modulation UWB system. The desired user's transmitted signal is [64]

$$s(t) = \sum_{k} p(t - kIT_f - d_k\delta t) \tag{2.165}$$

where k is the information bit index, I is the repetition length, T_f is the frame interval, $d_k \in \{0,1\}$ is the k-th information bit with equal *a priori* probabilities, δt is the additional time shift introduced when $d_k = 1$, and $p(t) = \sum_{n=0}^{I-1} g(t - nT_f - c_nT_c)$ is the symbol signal with $g(t)$ being the monocycle pulse with duration D_g, $\{c_n, 0 \le c_n \le N_h - 1\}$ being the time-hopping code and T_c being the chip time. After transmission over the UWB channel, the received signal can be expressed as [64]

$$y(t) = \sum_{l=1}^{L_c} \gamma_l s(t - \tau_l) + w(t) \tag{2.166}$$

where L_c is the number of multipath components assumed known, and γ_l and τ_l are the gain and the delay, respectively, of the l-th path to be estimated. The interference-plus-noise component is denoted $w(t)$ and is assumed white and Gaussian with power spectral density σ^2 [64].

The received signal is observed over $(0, T_0]$ where $T_0 = L \times (IT_f)$. As in [64], we assume that $\int_0^{T_0} s(t-\tau_{l_1})s(t-\tau_{l_2})dt \approx 0\ (l_1 \neq l_2)$ and $\tau_l < T_f - (N_h - 1)T_c - \delta\ (l = 1, 2, \cdots, L_c)$ such that the inter-path interference and the inter-frame interference can be ignored. For independent and identically distributed information bits, the log-likelihood function can be derived from (2.166) as

$$\ln\Lambda(\vartheta) = C - \frac{LE_p}{2\sigma^2}\sum_{l=1}^{L_c}\gamma_l^2 + \sum_{k=0}^{L-1}\left[\ln\cosh(a_k) + b_k\right] \tag{2.167}$$

where $\vartheta = [\gamma_1, \cdots, \gamma_{L_c}, \tau_1, \cdots, \tau_{L_c}]$ is the parameter vector, C is a constant independent of ϑ, $E_p = \int_0^{IT_f} p^2(t)dt$ is the signal energy, $\cosh(\cdot)$ is the hyperbolic cosine function, $a_k = \sum_{l=1}^{L_c}\gamma_l a_{kl}$, $a_{kl} = \frac{1}{2\sigma^2}\int_{t_k}^{t_k+IT_f} y(t)[p(t-t_k-\tau_l) - p(t-t_k-\tau_l-\delta t)]dt$, $b_k = \sum_{l=1}^{L_c}\gamma_l b_{kl}$, $b_{kl} = \frac{1}{2\sigma^2}\int_{t_k}^{t_k+IT_f} y(t)[p(t-t_k-\tau_l) + p(t-t_k-\tau_l-\delta t)]dt$, and $t_k = kIT_f$.

2.5.2 CRLB for NDA ML UWB Channel Estimation

From (2.167), the second order derivatives of the log-likelihood function with respect to ϑ can be derived as

$$\frac{\partial^2 \ln\Lambda(\vartheta)}{\partial\gamma_i\partial\gamma_j} = \sum_{k=0}^{L-1} a_{ki}a_{kj}\operatorname{sech}^2(a_k) \tag{2.168a}$$

$$\frac{\partial^2 \ln\Lambda(\vartheta)}{\partial\tau_i\partial\tau_j} = \gamma_i\gamma_j\sum_{k=0}^{L-1} a'_{ki}a'_{kj}\operatorname{sech}^2(a_k) \tag{2.168b}$$

$$\frac{\partial^2 \ln\Lambda(\vartheta)}{\partial\gamma_i\partial\tau_j} = \frac{\partial^2 \ln\Lambda(\vartheta)}{\partial\tau_j\partial\gamma_i} = \gamma_j\sum_{k=0}^{L-1} a_{ki}a'_{kj}\operatorname{sech}^2(a_k) \tag{2.168c}$$

for $i \neq j$ and $i, j = 1, 2, \cdots, L_c$, and

$$\frac{\partial^2 \ln\Lambda(\vartheta)}{\partial\gamma_i^2} = -\frac{LE_p}{\sigma^2} + \sum_{k=0}^{L-1} a_{ki}^2\operatorname{sech}^2(a_k) \tag{2.169a}$$

$$\frac{\partial^2 \ln\Lambda(\vartheta)}{\partial\tau_i^2} = \gamma_i\sum_{k=0}^{L-1}[b''_{ki} + b''_{ki}\tanh(a_k) + \gamma_i a'^2_{ki}\operatorname{sech}^2(a_k)] \tag{2.169b}$$

$$\frac{\partial^2 \ln\Lambda(\vartheta)}{\partial\gamma_i\partial\tau_i} = \frac{\partial^2 \ln\Lambda(\vartheta)}{\partial\tau_i\partial\gamma_i} = \sum_{k=0}^{L-1}[b'_{ki} + a'_{ki}\tanh(a_k) + \gamma_i a_{ki}a'_{ki}\operatorname{sech}^2(a_k)] \tag{2.169c}$$

for $i = 1, 2, \cdots, L_c$, where a'_{ki} and a''_{ki} are the first and the second order derivatives of a_{ki} with respect to τ_i, respectively, and b'_{ki} and b''_{ki} are the first and the second order derivatives of β_{ki} with respect to

τ_l, respectively. Denote

$$\Phi_2(\tau) = \int_0^{lT_f} p(t)p'(t+\tau)dt \tag{2.170a}$$

$$\Phi_3(\tau) = \int_0^{lT_f} p'(t)p'(t+\tau)dt \tag{2.170b}$$

$$\Phi_4(\tau) = \int_0^{lT_f} p(t)p''(t+\tau)dt. \tag{2.170c}$$

In a UWB system with large T_f, the functions in (2.170) are approximately zero when $D_g \leq |\tau| \leq T_f - (N_h - 1)T_c - \delta t$ (see, e.g., [66, Fig. 4]). Assume that the values of δt, $|\tau_{l_1} - \tau_{l_2}|$, and $|\tau_{l_1} - \tau_{l_2} \pm \delta t|$ are within this region for $l_1 \neq l_2$. It can be shown that a_{ki} are independent Gaussian random variables each with mean $(\tilde{d}_k\gamma_i\rho)$ and variance ρ, a'_{ki} are independent Gaussian random variables each with mean $(-\tilde{d}_k\gamma_i\rho\varepsilon)$ and variance $(\rho\eta)$, a''_{ki} are independent Gaussian random variables each with mean $(\tilde{d}_k\gamma_i\rho\zeta)$, b'_{ki} are independent Gaussian random variables each with mean $(-\gamma_i\rho\varepsilon)$, and b''_{ki} are independent Gaussian random variables each with mean $(\gamma_i\rho\zeta)$ for $k = 0, 1, \cdots, L-1$ and $i = 1, 2, \cdots, L_c$, where $\rho = E_p/(2\sigma^2)$ is the signal-to-interference-plus-noise ratio (SINR), $\varepsilon = \Phi_2(0)/E_p$, $\eta = \Phi_3(0)/E_p$, $\zeta = \Phi_4(0)/E_p$, $\tilde{d}_k = \text{sign}(1 - 2d_k)$, and $\text{sign}(\cdot)$ is the signum function. The values of ε, η and ζ are determined by the shape of $g(t)$. For a monocycle pulse satisfying $g^2(D_g) = g^2(0)$, one can further show that $\varepsilon = 0$ and a_{ki} is independent of a'_{ki}.

Using the PDFs of a_{ki}, a'_{ki}, a''_{ki}, b'_{ki} and b''_{ki} obtained, the expectations of the functions in (2.168) and (2.169) can be calculated. Let $R = \sum_{l=1}^{L_c} \gamma_l^2$. One has

$$-E\left\{\frac{\partial^2 \ln\Lambda(\vartheta)}{\partial\gamma_i\partial\gamma_j}\right\} = \begin{cases} -\frac{1}{\gamma_i\gamma_j}\sum_{k=0}^{L-1} F_{ijk} & i \neq j \\ 2L\rho - \frac{1}{\gamma_i^2}\sum_{k=0}^{L-1} G_{ik} & i = j \end{cases} \tag{2.171a}$$

$$-E\left\{\frac{\partial^2 \ln\Lambda(\vartheta)}{\partial\tau_i\partial\tau_j}\right\} = \begin{cases} 0 & i \neq j \\ -\sum_{k=0}^{L-1}[\rho\gamma_i^2 I_{ik} + H_{ik}] & i = j \end{cases} \tag{2.171b}$$

$$-E\left\{\frac{\partial^2 \ln\Lambda(\vartheta)}{\partial\gamma_i\partial\tau_j}\right\} = -E\left\{\frac{\partial^2 \ln\Lambda(\vartheta)}{\partial\tau_j\partial\gamma_i}\right\} = 0 \tag{2.171c}$$

where

$$F_{ijk} = \frac{1}{\sqrt{2\pi R\rho}}\int_{-\infty}^{\infty} \text{sech}^2(x) e^{-\frac{(x\tilde{d}_k - R\rho)^2}{2R\rho}} \cdot \left\{ [\rho\gamma_i^2 + \frac{\gamma_i^2(x\tilde{d}_k - R\rho + \rho\gamma_j^2)}{R - \gamma_j^2}][\rho\gamma_j^2 + \frac{\gamma_j^2(x\tilde{d}_k - R\rho)}{R}] \right.$$

$$-\frac{\gamma_i^2[\rho\gamma_j^2+\frac{\gamma_j^2}{R}(x\tilde{d}_k-R\rho)]^2}{R-\gamma_j^2}-\frac{\rho\gamma_i^2\gamma_j^2}{R}\Bigg\}dx \tag{2.172}$$

$$G_{ik}=\frac{1}{\sqrt{2\pi R\rho}}\int_{-\infty}^{\infty}\text{sech}^2(x)e^{-\frac{(x\tilde{d}_k-R\rho)^2}{2R\rho}}\cdot\left\{[\rho\gamma_i^2+\frac{\gamma_i^2(x\tilde{d}_k-R\rho)}{R}]^2+\frac{\rho(R-\gamma_i^2)\gamma_i^2}{R}\right\}dx \tag{2.173}$$

$$H_{ik} = \frac{\tilde{d}_k}{\sqrt{2\pi R\rho}}\int_{-\infty}^{\infty}\tanh(x)e^{-\frac{(x\tilde{d}_k-R\rho)^2}{2R\rho}}\cdot[\rho\zeta\gamma_i^2+\frac{\zeta\gamma_i^2}{R}(x\tilde{d}_k-R\rho)]dx \tag{2.174}$$

$$I_{ik}=\frac{\eta}{\sqrt{2\pi R\rho}}\int_{-\infty}^{\infty}\text{sech}^2(x)e^{-\frac{(x\tilde{d}_k-R\rho)^2}{2R\rho}}dx+\varepsilon. \tag{2.175}$$

The functions in (2.172)-(2.175) are obtained by taking the expectations of $\gamma_i\gamma_j a_{ki}a_{kj}\text{sech}^2(a_k)$, $\gamma_i^2 a_{ki}^2\text{sech}^2(a_k)$, $\gamma_i a_{ki}''\tanh(a_k)$, and $[b_{ki}''/\gamma_i+a'^2_{ki}\text{sech}^2(a_k)]/\rho$ with respect to the Gaussian random variables, respectively. They can be computed numerically.

Finally, denoting $\mathbf{A}(i,j)=-E\left\{\frac{\partial^2\ln\Lambda(\theta)}{\partial\gamma_i\partial\gamma_j}\right\}$ as the (i,j)-th element of the matrix $\mathbf{A}$, the CRLB for estimation of γ_l is $\mathbf{A}^{-1}(l,l)$, which is the l-th diagonal element of the inverse matrix $\mathbf{A}^{-1}$, and the CRLB for estimation of τ_l is $-1/\sum\limits_{k=0}^{L-1}[\rho\gamma_l^2 I_{lk}+H_{lk}]$. One sees that the CRLBs for estimation of γ_l and τ_l depend on d_k $(k=0,1,\cdots,L-1)$, γ_l $(l=1,2,\cdots,L_c)$, the SINR and the sample size.

Consider a sample size of $L=100$. Assume $d_k=0$ for $k=0,1,\cdots,99$. Similar to [64] and [65], the monocycle pulse is chosen as

$$g(t)=\left[1-16\pi\left(\frac{t-0.5D_g}{D_g}\right)^2\right]\exp[-8\pi(\frac{t-0.5D_g}{D_g})^2] \tag{2.176}$$

with $\varepsilon=0$, $\eta=125.61$ and $\zeta=-125.71$. Consider an UWB channel with $L_c=3$ and $R=1.1$. The SINR is $\rho=\frac{E_b}{2\sigma^2}$ with $E_b=E_p$. We denote the CRLB for NDA ML estimation of γ_1 as $CRLB_{\gamma-NDA}$, the CRLB for DA ML estimation of γ_1 derived in [65] as $CRLB_{\gamma-DA}$, and the corresponding CRLBs for ML estimation of τ_1 as $CRLB_{\tau-NDA}$ and $CRLB_{\tau-DA}$, respectively. Figs. 2.17 and 2.18 show the square roots of the CRLBs for ML estimation of γ_1 and τ_1, respectively, at different values of SINR. The multipath gains are fixed to $\gamma_1=0.73$, $\gamma_2=0.67$ and $\gamma_3=0.35$. One sees from Figs. 2.17 and 2.18 that $CRLB_{\gamma-NDA}$ and $CRLB_{\tau-NDA}$ are greater than $CRLB_{\gamma-DA}$ and $CRLB_{\tau-DA}$, respectively, when $\Gamma<7\ dB$, and they approach $CRLB_{\gamma-DA}$ and $CRLB_{\tau-DA}$, respectively, when

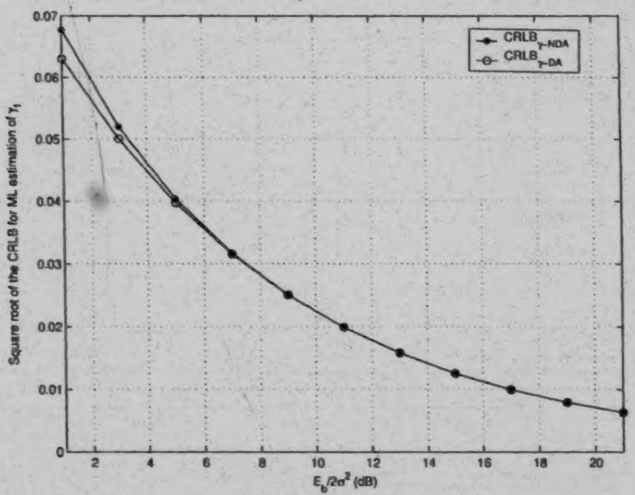

Figure 2.17. The square root of the CRLB for ML estimation of γ_1 at different values of ρ, when $\gamma_1 = 0.73$, $\gamma_2 = 0.67$ and $\gamma_3 = 0.35$. The CRLB for DA ML estimation is also shown.

$\Gamma > 7\ dB$. Figs. 2.19 and 2.20 show the square roots of the CRLBs for ML estimation of γ_1 and τ_1, respectively, at different values of γ_1. The SINR is set to $0\ dB$ and γ_3 is set to 0.35 in the calculations. Again, $CRLB_{\gamma-NDA}$ and $CRLB_{\tau-NDA}$ are greater than $CRLB_{\gamma-DA}$ and $CRLB_{\tau-DA}$, respectively. They decrease as γ_1 increases. Interestingly, the CRLB for the NDA case is close to the CRLB for the DA case, raising the question of whether one can obtain UWB channel estimation almost as good as in the DA case without the overhead expense of pilot symbols.

2.5.3 Novel NDA ML UWB Channel State Parameter Estimators

In [64], the authors approximated (2.167) with [64, eq. (25)]

$$\ln\Lambda(\vartheta) \approx -\frac{LE_p}{2\sigma^2}\sum_{l=1}^{L_c}\gamma_l^2 + \sum_{k=0}^{L-1} b_k \tag{2.177}$$

to derive NDA ML estimators for the UWB channel. The approximate log-likelihood function in (2.177) was obtained by using the approximation in [64, eq. (24)]. Equivalently, one can also obtain

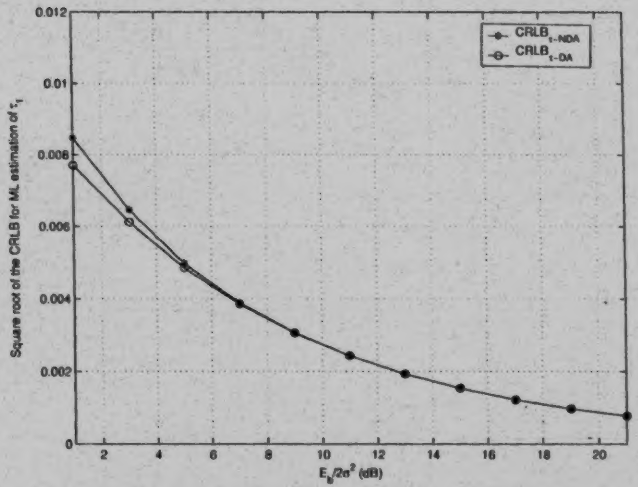

Figure 2.18. The square root of the CRLB for ML estimation of τ_1 at different values of ρ, when $\gamma_1 = 0.73$, $\gamma_2 = 0.67$ and $\gamma_3 = 0.35$. The CRLB for DA ML estimation is also shown.

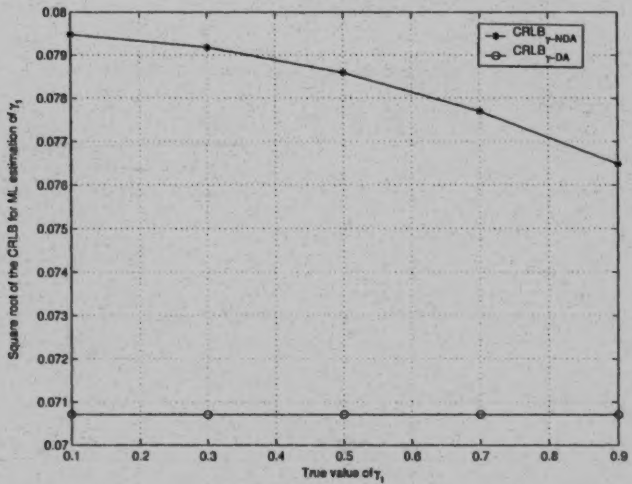

Figure 2.19. The square root of the CRLB for ML estimation of γ_1 at different values of γ_1, when $\rho = 0\ dB$. The CRLB for DA ML estimation is also shown.

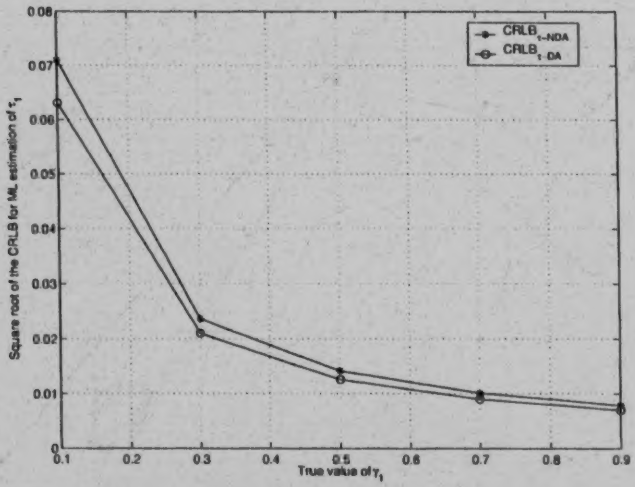

Figure 2.20. The square root of the CRLB for ML estimation of τ_1 at different values of γ_1, when $\rho = 0\,dB$. The CRLB for DA ML estimation is also shown.

(2.177) from (2.167) by using the approximation of

$$\ln\cosh(a_k) \approx 0. \tag{2.178}$$

Note that the performances of the NDA ML estimators depend on the accuracy of the approximation to (2.167). Note further that (2.178) has large approximation errors, especially when the SINR is large. One can improve the performances of the NDA ML estimators in [64] by using an approximation that is more accurate than (2.178). In this work, we propose using the approximation

$$\ln\cosh(a_k) \approx \sum_{l=1}^{L_c} \gamma_l |a_{kl}| \tag{2.179}$$

which is obtained by using the approximatiion $\ln\cosh(a_k) \approx |a_k|$ together with the approximation $|a_k| \approx \sum_{l=1}^{L_c} \gamma_l |a_{kl}|$. By using (2.179), the log-likelihood function in (2.167) can be approximated as

$$\ln\Lambda(\vartheta) \approx -\frac{LE_p}{2\sigma^2} \sum_{l=1}^{L_c} \gamma_l^2 + \sum_{k=0}^{L-1} \left[\sum_{l=1}^{L_c} \gamma_l |a_{kl}| + b_k \right]. \tag{2.180}$$

Comparing (2.180) with (2.177), one sees that there is an additional term $\sum_{k=0}^{L-1}\sum_{l=1}^{L_c}\gamma_l|a_{kl}|$ in (2.180). Similar to [64], one has from (2.180)

$$\ln\Lambda(\vartheta) \approx -\frac{LE_p}{2\sigma^2}\sum_{l=1}^{L_c}\left[\gamma_l - \frac{2\sigma^2}{LE_p}J(\tau_l)\right]^2 + \frac{2\sigma^2}{LE_p}J \qquad (2.181)$$

where $J(\tau_l) = \sum_{k=0}^{L-1}\frac{|a_{kl}|+b_{kl}}{2}$ and $J = \sum_{l=1}^{L_c} J^2(\tau_l)$. Since the maximum of J is found by maximizing each of $J^2(\tau_l), l = 1,2,\cdots,L_c$, the NDA ML estimates of $\tau_l, l = 1,2,\cdots,L_c$, are obtained by locating L_c maximas of $J^2(\tau)$ [64]. Denoting $\hat{\tau}_l$ as the estimate of τ_l, the NDA ML estimate of γ_l, $\hat{\gamma}_l$, is

$$\hat{\gamma}_l = \frac{2\sigma^2}{LE_p}|J(\hat{\tau}_l)| \qquad (2.182)$$

where we assume that the attenuation is always positive. The performances of the new NDA ML estimators, the NDA ML estimators and the DA ML estimators in [64] are compared as follows.

Without loss of generality, we focus on the estimation of τ_1 and γ_1. For convenience, we denote the new NDA ML estimators for τ_1 and γ_1 as τ_{NDA1} and γ_{NDA1}, respectively, the previous NDA ML estimators for τ_1 and γ_1 in [64] as τ_{NDA2} and γ_{NDA2}, respectively, and the previous DA ML estimators for τ_1 and γ_1 in [64] as τ_{DA} and γ_{DA}, respectively. Using the Gaussian monocycle as defined in (2.176), the received signal is sampled with a sampling interval of $T_s = 0.1D_g$. We choose $I = 5$, $N_h = 5$, $\delta t = 1.2D_g$, $T_f = 20D_g$, and $T_c = \frac{T_f}{N_h}$. Denote NU as the number of users in the UWB system. The cases when $NU = 1$, $NU = 10$ and $NU = 20$ are considered. We use a sample size of $L = 10$. Similar to [64], the number of multipath components is assumed to be $L_c = 3$, the multipath delays are the same for all users and are fixed at $\tau_l = 5lD_g$, $l = 1,2,3$, the multipath attenuations vary from user to user and are assumed to be independent Rayleigh random variables with an exponential power-delay profile of $E\{\gamma_l^2\} = De^{-l/4}$, $l = 1,2,3$, where D is the normalization factor. The desired user's multipath attenuations are fixed at $\gamma_1 = 0.73$, $\gamma_2 = 0.67$ and $\gamma_3 = 0.35$. The interfering users have transmitted signals similar to that in (2.165), except that their time origins are randomly selected from 0 to T_f to reflect the asynchronous operation. Each interfering user has the same transmitted signal power as the desired user. Note that, in the simulation, the values of T_f, I, N_h, L_c and L are fairly small, as the simulation time becomes intolerable for large values of T_f, I,

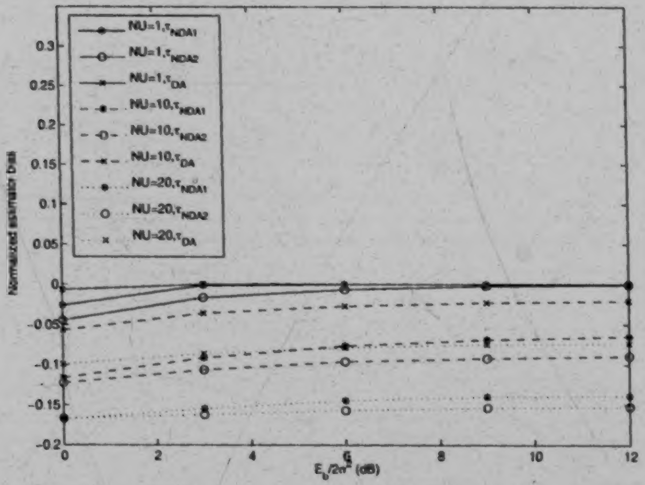

Figure 2.21. The normalized biases of τ_{NDA1}, τ_{NDA2} and τ_{DA} for different numbers of users.

N_h, L_c and L, and our purpose here is only to show that the new design leads to a better performance. In practice, the values of T_f, I, N_h, L_c and L will be much larger. Note also that, in addition to the MAI, the simulation carried out here also takes the inter-symbol and the inter-frame interference into account, as $\tau_l > T_f - (N_h - 1)T_c - \delta t$ for $l = 1,2,3$ in this case.

Figs. 2.21 and 2.22 show the normalized estimator biases and the normalized root mean squared errors (RMSEs) of τ_{NDA1}, τ_{NDA2} and τ_{DA}, respectively, where the normalization is with respect to D_g [64]. One sees that the performance of τ_{NDA1} improves when the SINR increases and/or the number of users decreases. At large values of SINR, the estimator exhibits a performance floor, caused mainly by the interferences in the system. Comparing τ_{NDA1} with τ_{NDA2}, one sees that τ_{NDA1} outperforms τ_{NDA2} for all the cases considered, as (2.179) has smaller approximation errors than (2.178). Also, comparing τ_{NDA1} with τ_{DA}, one observes that τ_{NDA1} underperforms τ_{DA}, as expected, as τ_{DA} uses pilot symbols. When $NU = 1$ and the SINR is large, τ_{NDA1} approaches τ_{DA}.

Figs. 2.23 and 2.24 show the normalized estimator biases and the normalized RMSEs of γ_{NDA1}, γ_{NDA2} and γ_{DA}, respectively, where the normalization is with respect to γ_1 [64]. Again, the performance of γ_{NDA1} improves as the SINR increases and/or the number of users decreases. Comparing

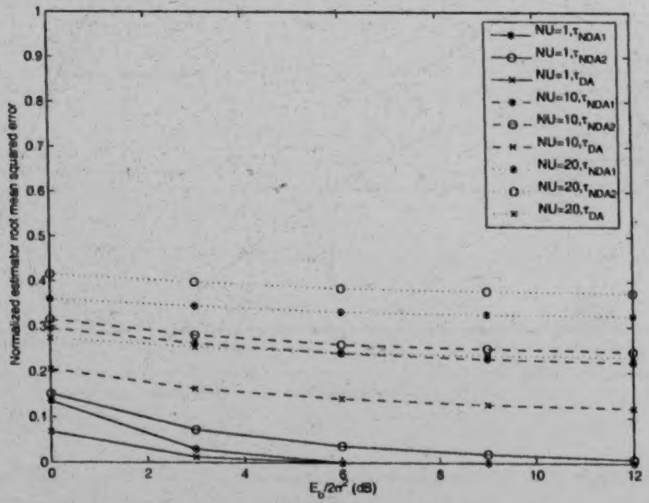

Figure 2.22. The normalized root mean squared errors of τ_{NDA1}, τ_{NDA2} and τ_{DA} for different numbers of users.

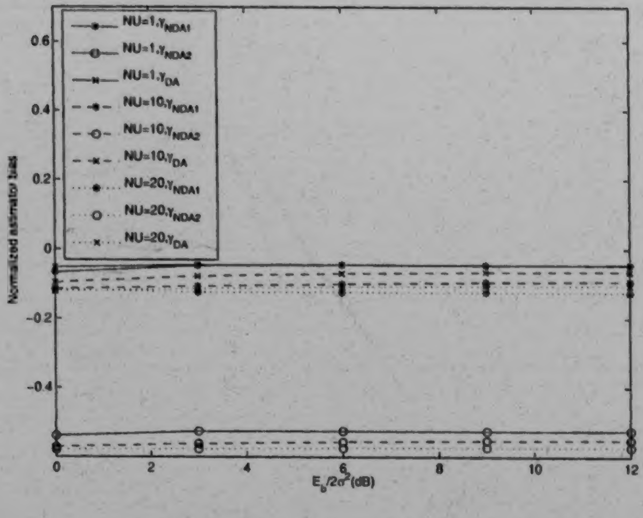

Figure 2.23. The normalized biases of γ_{NDA1}, γ_{NDA2} and γ_{DA} for different numbers of users.

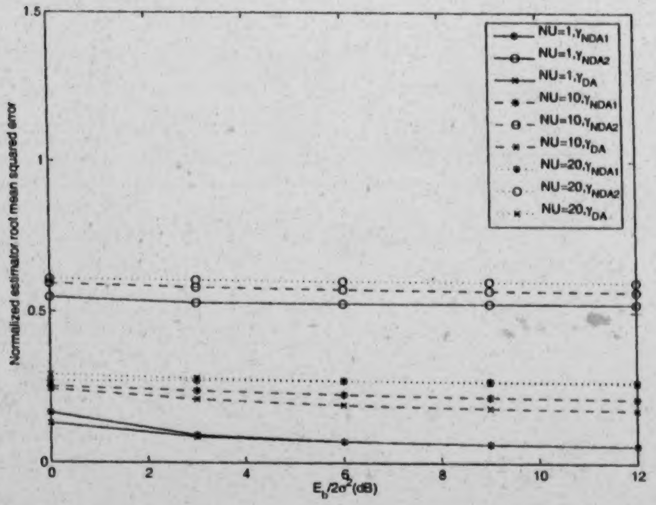

Figure 2.24. The normalized root mean squared errors of γ_{NDA1}, γ_{NDA2} and γ_{DA} for different numbers of users.

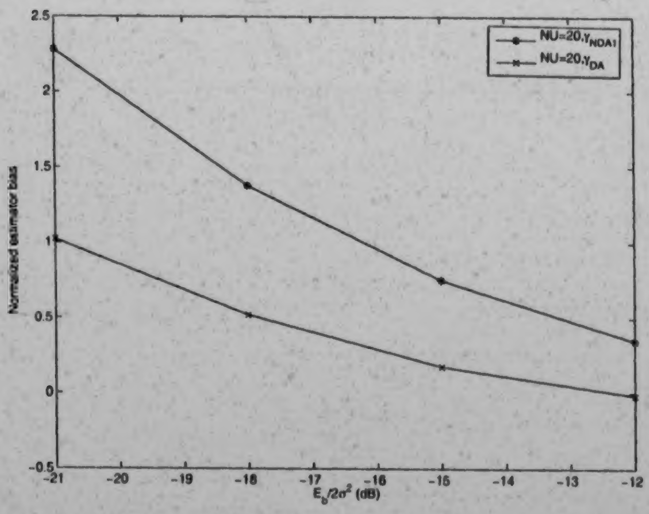

Figure 2.25. The normalized biases of γ_{NDA1} and γ_{DA} at small values of SINR.

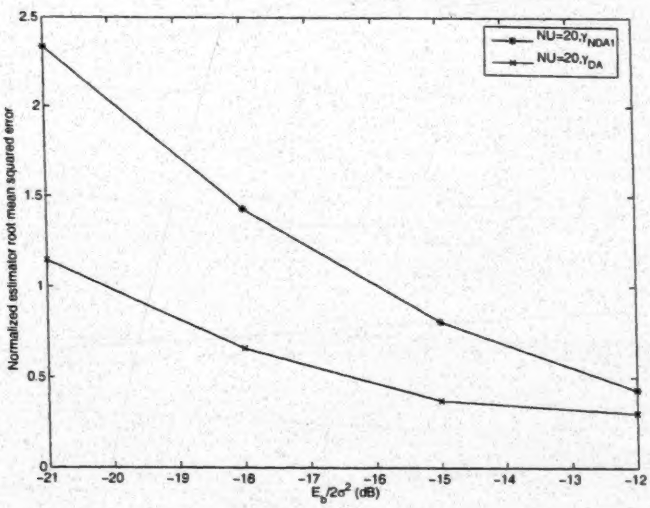

Figure 2.26. The normalized root mean squared errors of γ_{NDA1} and γ_{DA} at small values of SINR.

γ_{NDA1} with γ_{NDA2}, one sees that τ_{NDA1} performs at least 5 times better than τ_{NDA2} for all values of SINR considered. Also, comparing τ_{NDA1} with τ_{DA}, one sees that γ_{NDA1} underperforms γ_{DA} in most cases. However, when $NU = 20$ and the SINR is small, the RMSE of γ_{NDA1} is slightly smaller than that of γ_{DA}. This is caused by the fact that the estimate bias shifts from a negative value to a positive value as the SINR decreases. Figs. 2.25 and 2.26 compare the normalized estimator biases and the normalized RMSEs of γ_{NDA1} and γ_{DA} when the SINR decreases further. One sees that that the absolute value of the bias, and therefore, the RMSE of γ_{NDA1} are again larger than those of γ_{DA} when the SINR is small enough. This confirms our explanation. In the next chapter, we will study channel model parameter estimation in wireless communication channels.

Chapter 3

Wireless Channel Model Parameter Estimation

In the previous chapter, we have derived new channel state parameter estimators and new receivers that take channel estimation errors into account. Some applications may require knowledge of the channel model parameters as well. In this chapter, we estimate the model parameters in wireless communication channels. Specifically, we derive estimators for the parameters of three common fading distributions, Rayleigh, Ricean and Nakagami-m. Research in this area has been conducted by several researchers previously. In [9], [67]- [72], maximum likelihood (ML) and moment-based estimators for the Ricean parameters were proposed. In [73]- [80], estimation of the m parameter was studied in detail for the Nakagami-m fading channel. Some ML estimators for Ω on Rayleigh channels were derived in [81]. Most of these works are based on processing samples from a noiseless channel. However, in practical systems, one must process noisy samples. In this chapter, we first design new estimators for the fading distribution parameters by using noisy channel samples, with or without knowledge of the noise power. We then derive ML and approximate ML estimators for the Ricean K parameter by using fading phase samples, a method not studied previously.

3.1 Channel Model Parameter Estimation with Known Noise Power

In this section, we derive ML and/or moment-based estimators for Ω of the Rayleigh distribution in (1.1), K and Ω of the Ricean distribution in (1.3), and m and Ω of the Nakagami-m distribution in

(1.4), assuming known noise power.

3.1.1 System Model

The transmitted signal is assumed known in the estimation of the fading distribution parameters. This is the case for received signal samples taken during the transmission of a training sequence (which may be for channel estimation, synchronizer training or equalizer training). This is also the case when the receiver makes correct decisions which occur with high probability in a well designed system. The fading in the channel is assumed slow and flat. The fading signal is corrupted by additive white Gaussian noise (AWGN), which is independent of the fading. The received signal in the l-th symbol period can be expressed as

$$y_l(t) = r_l e^{j\theta_l} s_l(t) + n_l(t) \tag{3.1}$$

where $s_l(t)$ is the transmitted signal in the l-th symbol period, r_l is the fading envelope in the l-th symbol period having a Rayleigh, Ricean or Nakagami-m distribution, θ_l is the fading phase in the l-th symbol period and $n_l(t)$ is the complex AWGN on the channel in the l-th symbol interval with $E\{n_l(t)n_l^*(\tau)\} = N_0\delta(t-\tau)$ where $\delta(\cdot)$ is the impulse function. The value of N_0 is assumed known. The received signal $y_l(t)$ is correlated with the known transmitted signal $s_l(t)$ and normalized by its symbol energy E_l. The absolute value of the normalized correlator output is

$$z_l = |r_l e^{j\theta_l} + n_l| \tag{3.2}$$

where n_l is a complex Gaussian random variable with mean zero and variance $2\sigma_l^2 = \frac{N_0}{E_l}$. In the discussion, we assume that the transmitted signals are of equal energy. In this case, $E_l = E$ for any l and $2\sigma_l^2 = 2\sigma^2 = \frac{N_0}{E}$.

3.1.2 Estimators for the Rayleigh Distribution Parameter

When the channel is Rayleigh faded, it can be derived from (3.2) that the PDF of z_l is

$$f_{Z_l}(z_l) = \frac{2z_l}{\Omega + 2\sigma^2} e^{-\frac{z_l^2}{\Omega+2\sigma^2}}, \quad z_l \geq 0. \tag{3.3}$$

Using (3.3), the likelihood function for ML estimation of Ω can be derived. By differentiating the likelihood function with respect to Ω, setting the derivative equal to zero and solving the resulting equation, the ML estimator for Ω, $\hat{\Omega}_{Ray}$, is

$$\hat{\Omega}_{Ray} = \frac{1}{L}\sum_{l=1}^{L} z_l^2 - 2\sigma^2, \tag{3.4}$$

where L independent noisy channel samples are used. It can be proved that the likelihood function for $\Omega = 0$ is smaller than that for (3.4) when $\frac{1}{2\sigma^2 L}\sum_{l=1}^{L} z_l^2 > 1$. Note that when the noise level is small compared to the symbol energy, the term $2\sigma^2$ in (3.4) can be ignored and the resulting ML estimator for Ω is that derived in [81]. Note further that if the noiseless-based ML estimator derived in [81] is used in the noisy Rayleigh channel, the estimate obtained will be larger than that of our estimator in (3.4), giving a pessimistically modeled channel. Finally, if $\hat{\Omega}_{Ray} < 0$ in practice, the estimate will be discarded.

The mean and the variance of $\hat{\Omega}_{Ray}$ are derived as $E\{\hat{\Omega}_{Ray}\} = \Omega$ and $Var\{\hat{\Omega}_{Ray}\} = \frac{1}{L}(\Omega + 2\sigma^2)^2$. The Cramér-Rao lower bound (CRLB) for Ω can be derived analytically and is $\frac{1}{L}(\Omega + 2\sigma^2)^2$. Therefore, the ML estimator for Ω given in (3.4) is optimal in the sense of minimum variance.

3.1.3 Estimators for the Ricean Distribution Parameters

The PDF of z_l in a Ricean faded channel can be shown from (3.2) as

$$f_{z_l}(z_l) = \frac{z_l}{\sigma^2 + \alpha^2} e^{-\frac{z_l^2 + P^2}{2\sigma^2 + 2\alpha^2}} I_0\left(\frac{P z_l}{\sigma^2 + \alpha^2}\right), \quad z_l \geq 0, \tag{3.5}$$

which is again a Ricean distribution. Comparing (3.5) with (1.2), one notes that effectively the mean power of the line-of-sight (LOS) component remains the same while the mean power of the "scattering" components is changed from $2\alpha^2$ to $2\sigma^2 + 2\alpha^2$, showing the effect of noise in the estimation. Denoting $\Omega' = \Omega + 2\sigma^2 = P^2 + 2\alpha^2 + 2\sigma^2$ and $K' = \frac{P^2}{2\sigma^2 + 2\alpha^2}$, (3.5) can be rewritten in terms of Ω' and K'. The ML estimator for Ω', $\hat{\Omega}'$, is

$$\hat{\Omega}' = \frac{1}{L}\sum_{l=1}^{L} z_l^2 \tag{3.6}$$

and the ML estimator for K', $\hat{K}'$, is determined by the relation

$$\frac{1}{\hat{K}'+1} - 1 - \frac{1}{L}\sum_{l=1}^{L}\left(\frac{z_l}{\sqrt{\hat{\Omega}'}}\right)^2 + \frac{2\hat{K}'+1}{L\sqrt{\hat{K}'(\hat{K}'+1)}} \cdot \sum_{l=1}^{L} \frac{I_1\left(2\sqrt{\hat{K}'(\hat{K}'+1)}\frac{z_l}{\sqrt{\hat{\Omega}'}}\right)}{I_0\left(2\sqrt{\hat{K}'(\hat{K}'+1)}\frac{z_l}{\sqrt{\hat{\Omega}'}}\right)} \frac{z_l}{\sqrt{\hat{\Omega}'}} = 0 \tag{3.7}$$

where, again, L independent samples are used. The invariance principle of ML estimation maintains that the ML estimate of a transformation of several parameters equals the value obtained by applying the same transformation to ML estimates of each parameter [15]. Using the invariance principle here, one has that the ML estimators for Ω and K, $\hat{\Omega}_{Rice}$ and $\hat{K}$, are

$$\hat{\Omega}_{Rice} = \hat{\Omega}' - 2\sigma^2 \tag{3.8a}$$

$$\hat{K} = \frac{\hat{K}'\hat{\Omega}'}{\hat{\Omega}' - 2\sigma^2(\hat{K}'+1)}. \tag{3.8b}$$

where $\hat{\Omega}' - 2\sigma^2(\hat{K}'+1) > 0$ as $2\alpha^2\Omega' > 0$. Several observations can be made from the results in (3.6), (3.7) and (3.8). First, the ML estimates in (3.6) and (3.7) have the same forms as the noiseless-based ML estimators derived in reference [70]. But the ML estimators given in (3.6) and (3.7) use samples from a noisy Ricean channel while the noiseless-based ML estimators in reference [70] used samples from a noiseless Ricean channel. Second, the noiseless-sample-based estimators will pessimistically model the fading channel if they are used in a practical noisy channel. Finally, note that when the noise level in the channel is very small and/or the symbol energy is large, $2\sigma^2 \ll \hat{\Omega}'$. In this case, the noiseless-based ML estimators and the ML estimators derived here will yield nearly the same estimates.

The mean and the variance of $\hat{\Omega}_{Rice}$ are $E\{\hat{\Omega}_{Rice}\} = \Omega$ and $Var\{\hat{\Omega}_{Rice}\} = \frac{1}{L}(\frac{\Omega}{K+1} + 2\sigma^2)^2 + \frac{2K\Omega}{L(K+1)}(\frac{\Omega}{K+1} + 2\sigma^2)$. The CRLB for $\hat{\Omega}_{Rice}$ has to be calculated numerically. On the other hand, estimation of the K parameter requires deriving K' by solving a highly non-linear equation whose complexity increases with the sample size. Therefore, both performance analysis and implementation of (3.8b) are difficult. In the alternative, we propose some moment-based estimators for K. The n-th order moment of the noisy Ricean sample is [1, p. 48]

$$\mu_n = E\{z_l^n\} = \left(\frac{\Omega'}{K'+1}\right)^{\frac{n}{2}}\Gamma\left(\frac{n}{2}+1\right)e^{-K'}\,{}_1F_1\left(\frac{n}{2}+1, 1; K'\right) \tag{3.9}$$

where ${}_1F_1(\cdot,\cdot;\cdot)$ is the confluent hypergeometric function [2, p. 504]. As we mentioned in Chapter 1, in moment-based estimation, the true value of μ_n is usually approximated by the value of $\hat{\mu}_n = \frac{1}{L}\sum_{l=1}^{L} z_l^n$. By extending the result of [68, eqn. (3)] to the noisy sample case, one moment-based estimator for K', $\hat{K'}_{MB1}$, is

$$\hat{K'}_{MB1} = \frac{\sqrt{2-\frac{\hat{\mu}_4}{\hat{\mu}_2^2}}}{1-\sqrt{2-\frac{\hat{\mu}_4}{\hat{\mu}_2^2}}}. \tag{3.10}$$

The corresponding moment-based estimator for K, $\hat{K}_{MB1}$, can be derived by using (3.10) in (3.8b), resulting in the estimator

$$\hat{K}_{MB1} = \frac{\hat{\mu}_2\sqrt{2-\frac{\hat{\mu}_4}{\hat{\mu}_2^2}}}{\hat{\mu}_2\left(1-\sqrt{2-\frac{\hat{\mu}_4}{\hat{\mu}_2^2}}\right)-2\sigma^2} \tag{3.11}$$

where the ML estimator for $\hat{\Omega}'$ given in (3.6) is used.

Another moment-based estimator can be derived by using the relation $a\,{}_1F_1(a+1,b;x) = (x+2a-b)\,{}_1F_1(a,b;x)+(b-a)\,{}_1F_1(a-1,b;x)$ [2, eqn. (13.4.1)]. One has the moment-based estimator for K', $\hat{K'}_{MB2}$, as

$$\hat{K'}_{MB2} = \frac{3\hat{\Omega}'\hat{\mu}_1 - 2\hat{\mu}_3 + \hat{\Omega}'\sqrt{\hat{\mu}_1^2 - \hat{\mu}_{-1}\hat{\mu}_3 + \hat{\mu}_{-1}\hat{\mu}_1\hat{\Omega}'}}{2\hat{\mu}_3 - 2\hat{\mu}_1\hat{\Omega}'} \tag{3.12}$$

since $\frac{3\mu_3}{2\Gamma(5/2)} = \frac{(K'+2)\Omega'\mu_1}{(K'+1)\Gamma(3/2)} - \frac{\Omega'^2\mu_{-1}}{2\Gamma(1/2)(K'+1)^2}$. The estimator for K, $\hat{K}_{MB2}$, is derived by using (3.12) in (3.8b) as

$$\hat{K}_{MB2} = \frac{\hat{\mu}_2(3\hat{\mu}_2\hat{\mu}_1 - 2\hat{\mu}_3 + a)}{\hat{\mu}_2(2\hat{\mu}_3 - 2\hat{\mu}_1\hat{\mu}_2) - 2\sigma^2(\hat{\mu}_2\hat{\mu}_1 + a)} \tag{3.13}$$

where $a = \hat{\mu}_2\sqrt{\hat{\mu}_1^2 - \hat{\mu}_{-1}\hat{\mu}_3 + \hat{\mu}_{-1}\hat{\mu}_1\hat{\mu}_2}$ and the ML estimator for Ω' given in (3.6) is also used.

3.1.4 Estimators for the Nakagami-*m* Distribution Parameters

As previously, one has to find the PDF of z_l in order to derive estimators for the parameters m and Ω. It can be shown by integration that

$$f_{Z_l}(z_l) \quad = \quad \frac{z_l}{\sigma^2(\frac{\Omega}{2\sigma^2 m}+1)^m} e^{-\frac{z_l^2}{2\sigma^2}}\,{}_1F_1(m,1;\frac{z_l^2}{2\sigma^2+\frac{4\sigma^4 m}{\Omega}}), \quad z_l \geq 0 \tag{3.14}$$

in a noisy Nakagami-m fading channel. Note that if the asymptotic expansion of the confluent hypergeometric function [2, eqn. (13.5.1)] is used in (3.14) and $2\sigma^2$ is set equal to 0, the PDF of the noiseless samples in a Nakagami-m fading channel can be derived, as expected. Using (3.14), the likelihood function can be derived and the ML estimators for m and Ω could be found by differentiating it with respect to m and Ω. Unfortunately, the results are complex and do not lead to tractable estimators. We propose moment-based estimators for Ω and m instead. The n-th order moment of the noisy Nakagami-m channel sample is

$$\mu_n = E\{z_l^n\} = (2\sigma^2)^{\frac{n}{2}}\Gamma(\frac{n}{2}+1)(\frac{2\sigma^2 m}{\Omega+2\sigma^2 m})^m \cdot F(m,\frac{n}{2}+1;1;\frac{\Omega}{\Omega+2\sigma^2 m}) \quad (3.15)$$

where $F(\cdot,\cdot;\cdot;\cdot)$ is the hypergeometric function [2, p. 556]. From [2, eqn. (15.2.11)], one has $(\mu_2 - 2\sigma^2)\Omega + (6\sigma^2\mu_2 - 4\sigma^4 - \mu_4)m + \mu_2\Omega m = 0$ and $(\mu_1 - \sigma^2\mu_{-1})\Omega + (8\sigma^2\mu_1 - 2\mu_3 - 2\sigma^4\mu_{-1})m + 2\mu_1 m\Omega = 0$. Solving the equations for m and Ω, moment-based estimators for m and Ω are

$$\hat{m} = \frac{a_2(b_1c_2 - b_2c_1) + b_2(a_2c_1 - a_1c_2)}{c_2(b_2c_1 - b_1c_2)} \quad (3.16)$$

$$\hat{\Omega}_{Naka} = \frac{a_2(b_1c_2 - b_2c_1) + b_2(a_2c_1 - a_1c_2)}{c_2(a_1c_2 - a_2c_1)} \quad (3.17)$$

where $a_1 = \hat{\mu}_2 - 2\sigma^2$, $b_1 = 6\sigma^2\hat{\mu}_2 - 4\sigma^4 - \hat{\mu}_4$, $c_1 = \hat{\mu}_2$, $a_2 = \hat{\mu}_1 - \sigma^2\hat{\mu}_{-1}$, $b_2 = 8\sigma^2\hat{\mu}_1 - 2\hat{\mu}_3 - 2\sigma^4\hat{\mu}_{-1}$ and $c_2 = 2\hat{\mu}_1$.

3.1.5 Numerical Results

Here, we use Monte Carlo simulation to examine the fading distribution parameter estimators in terms of the sample means and sample root mean squared errors (RMSEs). Sample sizes of $L = 100$, $L = 500$ or $L = 1000$ are used and noise variances of $\sigma^2 = 0.1$ and $\sigma^2 = 0.5$ are considered. Previous work [68]- [80] has shown that fewer samples do not provide reliable estimates.

Figs. 3.1 to 3.3 show the performances of the moment-based estimatiors $\hat{K}_{MB1}$ given in (3.11) and $\hat{K}_{MB2}$ given in (3.13), where negative estimates have been set equal to 0 when obtained. Figs. 3.1 and 3.2 show the estimator means normalized to the true value of K to better show the differences. The estimators $\hat{K}_{MB1}$ and $\hat{K}_{MB2}$ have positive biases, but the biases are very small, less than 3% for

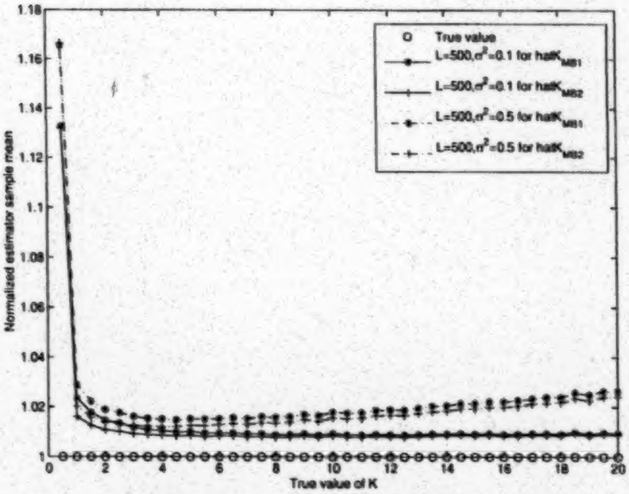

Figure 3.1. Comparison of the normalized sample means of $\hat{K}_{MB1}$ given in (3.11) and $\hat{K}_{MB2}$ given in (3.13) with the true value for $L = 500$, $\sigma^2 = 0.1$ and $L = 500$, $\sigma^2 = 0.5$.

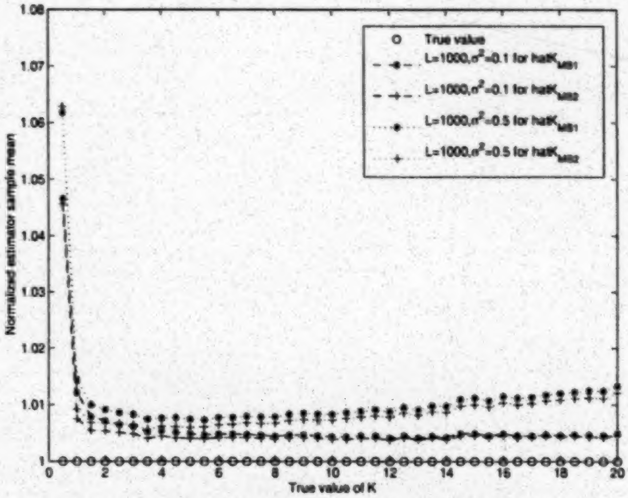

Figure 3.2. Comparison of the normalized sample means of $\hat{K}_{MB1}$ given in (3.11) and $\hat{K}_{MB2}$ given in (3.13) with the true value for $L = 1000$, $\sigma^2 = 0.1$ and $L = 1000$, $\sigma^2 = 0.5$.

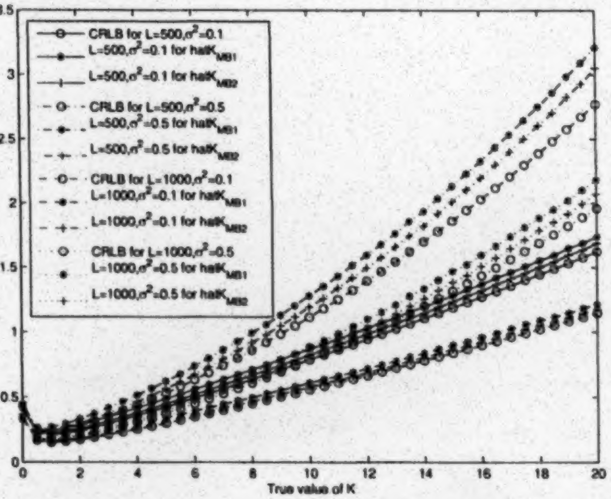

Figure 3.3. Comparison of the RMSEs of $\hat{K}_{MB1}$ given in (3.11) and $\hat{K}_{MB2}$ given in (3.13) with the CRLBs for $L = 500$, $\sigma^2 = 0.1$, $L = 500$, $\sigma^2 = 0.5$, $L = 1000$, $\sigma^2 = 0.1$ and $L = 1000$, $\sigma^2 = 0.5$.

most values of K shown in Figs. 3.1 and 3.2. Also, in the case when $\sigma^2 = 0.5$, the normalized mean increases with the value of K, while in the case when $\sigma^2 = 0.1$, the normalized mean nearly remains constant. Fig. 3.3 shows the RMSEs of the estimators. Also shown in Fig. 3.3 are the corresponding CRLBs, computed numerically, to benchmark the RMSEs of the estimators. The CRLBs for $K = 0$ go to infinity, and they are not shown in the figures. These estimators have small positive deviations from the CRLBs. The biases and the deviations can be reduced by increasing the sample size and/or decreasing the noise variance, as can be seen in Figs. 3.1 to 3.3. Comparing $\hat{K}_{MB1}$ with $\hat{K}_{MB2}$, one notes that $\hat{K}_{MB2}$ outperforms $\hat{K}_{MB1}$. This is expected because the former uses lower order moments and, therefore, is less susceptible to large outlying noise samples. Figs. 3.4 and 3.5 show the performance of $\hat{\Omega}_{Rice}$ given in (3.8a). The bias of $\hat{\Omega}_{Rice}$ is less than 0.1% of the true value. It attains the CRLBs graphically when $L = 1000$. When $L = 100$, the estimator has very small deviations from the CRLBs at large values of Ω. Figs. 3.6 and 3.7 show the performance of $\hat{\Omega}_{Ray}$ given in (3.4). One sees that it has similar performance to $\hat{\Omega}_{Rice}$.

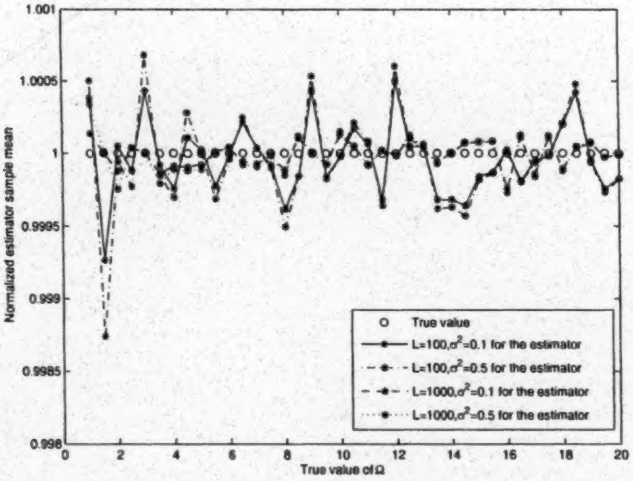

Figure 3.4. Comparison of the normalized sample mean of $\hat{\Omega}_{Rice}$ given in (3.8a) with the true value for $L = 100$, $\sigma^2 = 0.1$, $L = 100$, $\sigma^2 = 0.5$, $L = 1000$, $\sigma^2 = 0.1$ and $L = 1000$, $\sigma^2 = 0.5$.

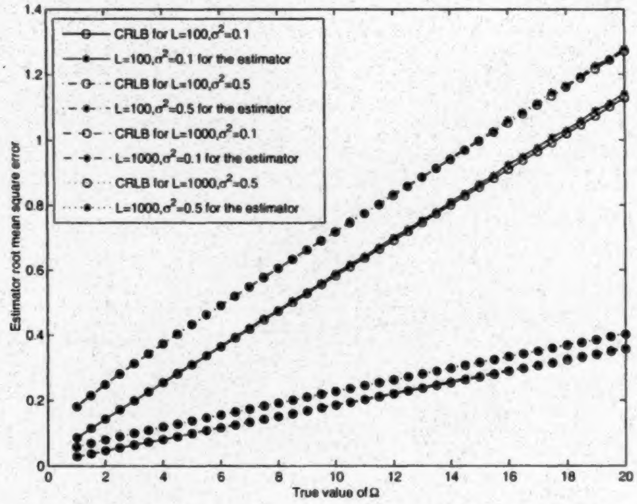

Figure 3.5. Comparison of the RMSE of $\hat{\Omega}_{Rice}$ given in (3.8a) with the CRLBs for $L = 100$, $\sigma^2 = 0.1$, $L = 100$, $\sigma^2 = 0.5$, $L = 1000$, $\sigma^2 = 0.1$ and $L = 1000$, $\sigma^2 = 0.5$.

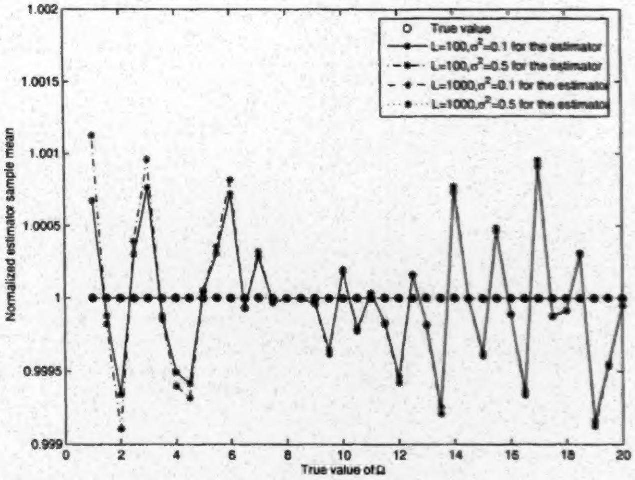

Figure 3.6. Comparison of the normalized sample mean of $\hat{\Omega}_{Ray}$ given in (3.4) with the true value for $L = 100$, $\sigma^2 = 0.1$, $L = 100$, $\sigma^2 = 0.5$, $L = 1000$, $\sigma^2 = 0.1$ and $L = 1000$, $\sigma^2 = 0.5$.

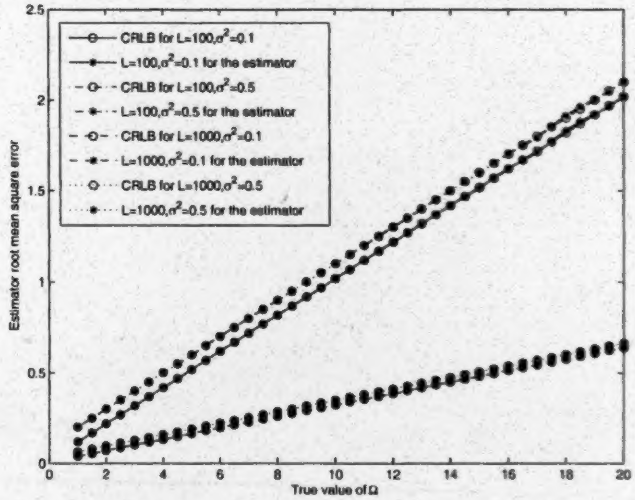

Figure 3.7. Comparison of the RMSE of $\hat{\Omega}_{Ray}$ given in (3.4) with the CRLBs for $L = 100$, $\sigma^2 = 0.1$, $L = 100$, $\sigma^2 = 0.5$, $L = 1000$, $\sigma^2 = 0.1$ and $L = 1000$, $\sigma^2 = 0.5$.

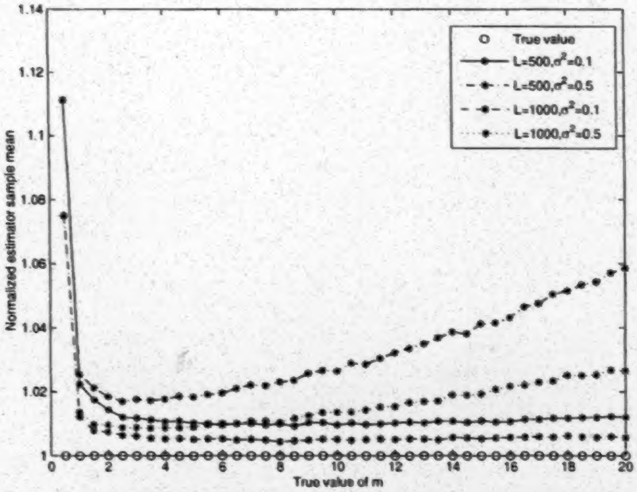

Figure 3.8. Comparison of the normalized sample mean of $\hat{m}$ given in (3.16) with the true value for $L = 500$, $\sigma^2 = 0.1$, $L = 500$, $\sigma^2 = 0.5$, $L = 1000$, $\sigma^2 = 0.1$ and $L = 1000$, $\sigma^2 = 0.5$.

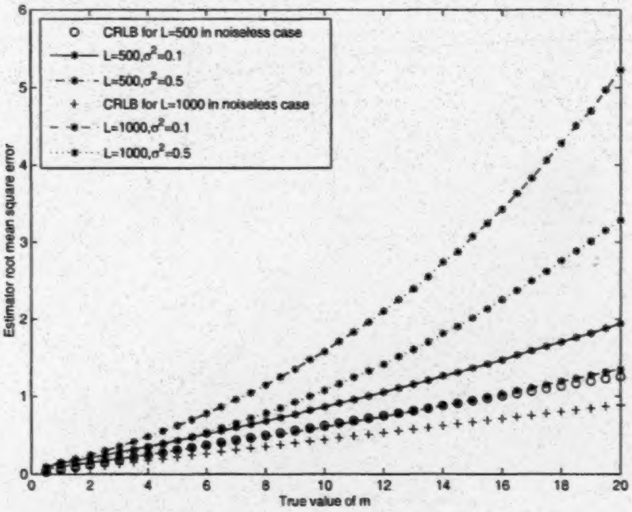

Figure 3.9. Comparison of the RMSE of $\hat{m}$ given in (3.16) with the CRLBs for $L = 500$, $\sigma^2 = 0.1$, $L = 500$, $\sigma^2 = 0.5$, $L = 1000$, $\sigma^2 = 0.1$ and $L = 1000$, $\sigma^2 = 0.5$.

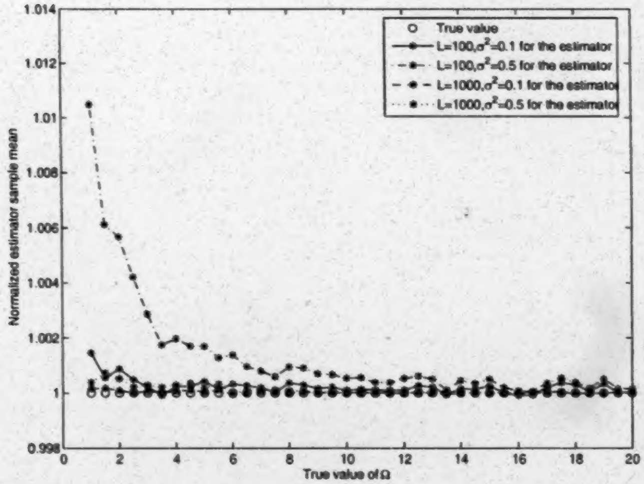

Figure 3.10. Comparison of the normalized sample mean of $\hat{\Omega}_{Naka}$ given in (3.17) with the true value for $L = 100$, $\sigma^2 = 0.1$, $L = 100$, $\sigma^2 = 0.5$, $L = 1000$, $\sigma^2 = 0.1$ and $L = 1000$, $\sigma^2 = 0.5$.

Figs. 3.8 and 3.9 show the normalized sample mean and RMSE, respectively, of $\hat{m}$ given in (3.16), where estimates less than 0.5 have been set equal to 0.5 when obtained. The derivation of the CRLB for m is untractable in a noisy Nakagami-m channel. The CRLB for m in the noiseless Nakagami-m channel is used instead in Fig. 3.9. One sees from Fig. 3.8 that the estimator $\hat{m}$ has a positive bias, but the bias is small, less than 5% for most values of m. Also, from Fig. 3.8, one sees that the noise has a greater influence on the estimator performance than does the sample size. Consider the case when $\sigma^2 = 0.5$; not only is the normalized mean in this case larger than that in the case when $\sigma^2 = 0.1$, but the normalized mean increases with the value of m as well. From Fig. 3.9, one sees that the estimator $\hat{m}$ has a positive deviation from the noiseless CRLB. The performance of the estimator $\hat{m}$ is greatly improved when the noise variance decreases and/or the sample size increases, as can be seen in Figs. 3.8 and 3.9. Figs. 3.10 and 3.11 show the performance of $\hat{\Omega}_{Naka}$ given in (3.17). Most of the biases of $\hat{\Omega}_{Naka}$ are less than 1%. Also, even when the sample size is small (L=100) and the noise variance is high ($\sigma^2 = 0.5$), the deviation from the CRLB is still small.

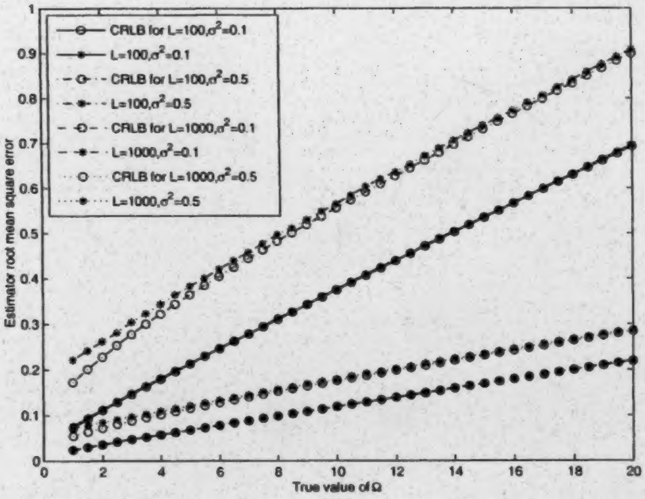

Figure 3.11. Comparison of the RMSE of $\hat{\Omega}_{Naka}$ given in (3.17) with the CRLBs for $L = 100$, $\sigma^2 = 0.1$, $L = 100$, $\sigma^2 = 0.5$, $L = 1000$, $\sigma^2 = 0.1$ and $L = 1000$, $\sigma^2 = 0.5$.

3.2 Channel Model Parameter Estimation with Unknown Noise Power

In the previous section, novel noise-based estimators for fading distribution parameters were developed. Derivation of these estimators requires knowledge of the noise power in the channel. However, in some cases, it is necessary to estimate fading distribution parameters without knowledge of the noise power in the channel, and therefore, it is required to jointly estimate the noise power and the fading distribution parameters. In this work, estimation of fading distribution parameters without knowledge of the noise power in the channel is studied.

3.2.1 System Model

Use similar assumptions to those previously, except that the values of N_0 or $2\sigma^2$ are unknown now. Express the received signal in (3.1) in a vector form, where the i-th component is

$$y_{il} = r_l e^{j\theta_l} s_{il} + n_{il} \tag{3.18}$$

with $y_{il} = \int_T y_l(t) f_i(t) dt$, $s_{il} = \int_T s_l(t) f_i(t) dt$, $n_{il} = \int_T n_l(t) f_i(t) dt$, T being the time duration of one symbol signal, and $\{f_i(t)\}_{i=1}^{I}$ being a set of I orthonormal functions. The noise samples n_{il}, $i = 1, 2, \cdots, I$, are independent Gaussian random variables each with mean zero and variance N_0 [1]. The likelihood function in terms of the samples obtained from the l-th symbol can be derived as

$$f(\mathbf{y}_l | N_0, r_l, \theta_l) = \left(\frac{1}{\pi N_0} \right)^I \exp\left\{ -\frac{1}{N_0} \sum_{i=1}^{I} |y_{il} - r_l e^{j\theta_l} s_{il}|^2 \right\} \tag{3.19}$$

where $\mathbf{y}_l = [y_{1l}, y_{2l}, \cdots, y_{Il}]$ is the received signal samples in the l-th symbol interval. Denote

$$E_l = \sum_{i=1}^{I} |s_{il}|^2 \tag{3.20}$$

$$z_l = \frac{1}{E_l} \left| \sum_{i=1}^{I} y_{il} s_{il}^* \right| \tag{3.21}$$

and

$$\psi_l = \arctan \left[\frac{Im\{\sum_{i=1}^{I} y_{il} s_{il}^*\}}{Re\{\sum_{i=1}^{I} y_{il} s_{il}^*\}} \right] \tag{3.22}$$

as the energy of the l-th transmitted signal, the envelope and the phase of the normalized correlator output correlating the received signal and the transmitted signal in the l-th symbol interval, respectively. Note that (3.21) is also given in (3.2). As previously, we consider the case when the energies of the transmitted signals are the same. In this case, $E_l = E$ and $2\sigma_l^2 = 2\sigma^2 = \frac{N_0}{E}$ for all l. Also, define $\frac{E}{N_0}$ as the transmitted-signal-to-noise ratio (TSNR) (the average signal-to-noise ratio in the fading channel equals the product of the TSNR and the mean fading power). The likelihood function in (3.19) can be rewritten as

$$f(\mathbf{y}_l | \sigma, r_l, \theta_l) = \left(\frac{1}{2\pi E \sigma^2} \right)^I \exp\left\{ -\frac{1}{2E\sigma^2} \sum_{i=1}^{I} |y_{il}|^2 - \frac{r_l^2}{2\sigma^2} + \frac{r_l}{\sigma^2} z_l \cos(\theta_l - \psi_l) \right\} \tag{3.23}$$

where $2\sigma^2$ is unknown.

3.2.2 Estimators for the Rayleigh Distribution Parameter

In order to estimate Ω, a joint estimation of $2\sigma^2$ and Ω is needed as the noise power in the channel is also unknown. The likelihood function for joint ML estimation of $2\sigma^2$ and Ω in terms of the samples obtained from the l-th symbol can be derived as

$$f(\mathbf{y}_l|\sigma,\Omega) = \int_0^\infty \int_{-\pi}^{\pi} f(\mathbf{y}_l|\sigma,r_l,\theta_l) f_R(r_l|\Omega) f_\Theta(\theta_l) dr_l d\theta_l \tag{3.24}$$

where $f(\mathbf{y}_l|\sigma,r_l,\theta_l)$ is the likelihood function given by (3.23), $f_R(r_l|\Omega)$ is the PDF of the Rayleigh distribution given by (1.1), and $f_\Theta(\theta_l) = \frac{1}{2\pi}$ is the PDF of the fading phase. Solving the integral in (3.24), one has

$$f(\mathbf{y}_l|\sigma,\Omega) = \left(\frac{1}{2\pi E\sigma^2}\right)^I \frac{2\sigma^2}{2\sigma^2+\Omega} \exp\left\{-\frac{1}{2E\sigma^2}\sum_{i=1}^{I}|y_{il}|^2 + \frac{z_l^2}{2\sigma^2} - \frac{z_l^2}{2\sigma^2+\Omega}\right\}. \tag{3.25}$$

In practice, more accurate estimate can be obtained by using a sequence of symbols instead of one single symbol. Assuming L independent symbols are available, the likelihood function in terms of the samples obtained from the L symbols can be derived from (3.25) as

$$f(\mathbf{y}|\sigma,\Omega) = \left(\frac{1}{2\pi E\sigma^2}\right)^{LI} \left(\frac{2\sigma^2}{2\sigma^2+\Omega}\right)^L \cdot \exp\left\{-\frac{\sum_{l=1}^{L}\sum_{i=1}^{I}|y_{il}|^2}{2E\sigma^2} + \frac{\sum_{l=1}^{L} z_l^2}{2\sigma^2} - \frac{\sum_{l=1}^{L} z_l^2}{2\sigma^2+\Omega}\right\} \tag{3.26}$$

where $\mathbf{y} = [\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_L]$ is the sample vector of L symbols. The ML estimators for $2\sigma^2$ and Ω in a noisy Rayleigh fading channel can be derived as

$$\hat{2\sigma^2} = \frac{1}{(I-1)LE}\sum_{l=1}^{L}\sum_{i=1}^{I}|y_{il}|^2 - \frac{1}{(I-1)L}\sum_{l=1}^{L} z_l^2 \tag{3.27}$$

and

$$\hat{\Omega}_{Ray} = \hat{\Omega}'_{Ray} - \hat{2\sigma^2} \tag{3.28}$$

where $\hat{\Omega}'_{Ray} = \frac{1}{L}\sum_{l=1}^{L} z_l^2$. Note that the ML estimator for Ω in (3.28) derived with unknown $2\sigma^2$ has similar form to that in (3.4) derived with known $2\sigma^2$, except that the true value of $2\sigma^2$ in (3.4) is replaced by the ML estimate of $2\sigma^2$ given in (3.27).

The performances of the estimators $\hat{2\sigma^2}$ and $\hat{\Omega}_{Ray}$ can be analyzed as follows. Rewrite (3.27) as

$$\hat{2\sigma^2} = \frac{1}{(I-1)LE}\sum_{l=1}^{L}\left[\sum_{i=1}^{I}|y_{il}|^2 - Ez_l^2\right]. \tag{3.29}$$

It is proved in Appendix B that $\sum_{i=1}^{I}|y_{il}|^2 - Ez_l^2$ is independent of z_l and has a central chi-square distribution with $(2I-2)$ degrees of freedom and parameter $\frac{N_0}{2}$. Moreover, since noisy samples taken during different symbol intervals are independent, $\sum_{l=1}^{L}\left[\sum_{i=1}^{I}|y_{il}|^2 - Ez_l^2\right]$ is also independent of $\sum_{l=1}^{L} z_l^2$. Thus, $2\hat{\sigma^2}$ is indpendent of $\hat{\Omega}'_{Ray}$ and has a central chi-square distribution with $(2IL-2L)$ degrees of freedom and parameter $\frac{N_0}{2(I-1)LE}$. The mean and the variance of $2\hat{\sigma^2}$ are $E\{2\hat{\sigma^2}\} = 2\sigma^2$ and $Var\{2\hat{\sigma^2}\} = \frac{1}{(I-1)L}\left(2\sigma^2\right)^2$. The performance of $\hat{\Omega}_{Ray}$ in (3.28) can be determined similarly. Using results in the previous section, one can derive $E\{z_l^2\} = \Omega + 2\sigma^2$ and $Var\{z_l^2\} = (\Omega+2\sigma^2)^2$. Since the random variables $\hat{\Omega}'_{Ray}$ and $2\hat{\sigma^2}$ in (3.28) are independent, the mean and the variance of $\hat{\Omega}_{Ray}$ are $E\{\hat{\Omega}_{Ray}\} = \Omega$ and $Var\{\hat{\Omega}_{Ray}\} = \frac{1}{L}(\Omega+2\sigma^2)^2 + \frac{1}{(I-1)L}\left(2\sigma^2\right)^2$.

A benchmark for unbiased estimators is the Cramér-Rao lower bound (CRLB) [17]. Since both $2\sigma^2$ and Ω are unknown, the CRLBs for estimation of $2\sigma^2$ and Ω in a noisy Rayleigh fading channel are the (1,1) and (2,2) elements of $\mathbf{J}^{-1}$ [17, eqn. (259)], respectively, where $\mathbf{J}$ is the Fisher information matrix defined as

$$\mathbf{J} = \begin{bmatrix} -E\{\frac{\partial^2 \ln f}{\partial(2\sigma^2)^2}\} & -E\{\frac{\partial^2 \ln f}{\partial 2\sigma^2 \partial\Omega}\} \\ -E\{\frac{\partial^2 \ln f}{\partial\Omega\partial 2\sigma^2}\} & -E\{\frac{\partial^2 \ln f}{\partial\Omega^2}\} \end{bmatrix}. \tag{3.30}$$

Using the likelihood function in (3.26), the elements in $\mathbf{J}$ can be derived as $-E\{\frac{\partial^2 \ln f}{\partial(2\sigma^2)^2}\} = \frac{(I-1)L}{(2\sigma^2)^2} + \frac{L}{(2\sigma^2+\Omega)^2}$, $-E\{\frac{\partial^2 \ln f}{\partial\Omega^2}\} = \frac{L}{(2\sigma^2+\Omega)^2}$, and $-E\{\frac{\partial^2 \ln f}{\partial 2\sigma^2\partial\Omega}\} = -E\{\frac{\partial^2 \ln f}{\partial\Omega\partial 2\sigma^2}\} = \frac{L}{(2\sigma^2+\Omega)^2}$. Putting these in (3.30) and solving (3.30) for its inverse matrix, the CRLB for estimation of $2\sigma^2$ in a noisy Rayleigh fading channel is $\frac{1}{(I-1)L}\left(2\sigma^2\right)^2$ and the CRLB for estimation of Ω in this case is $\frac{1}{L}(\Omega+2\sigma^2)^2 + \frac{1}{(I-1)L}\left(2\sigma^2\right)^2$. Therefore, the ML estimators for $2\sigma^2$ and Ω in a noisy Rayleigh fading channel, $2\hat{\sigma^2}$ and $\hat{\Omega}_{Ray}$, given by (3.27) and (3.28), respectively, are unbiased and achieve the CRLB's. They are optimal in the sense of minimum variance.

3.2.3 Estimators for the Ricean Distribution Parameters

Here, we derive estimators for Ricean distribution parameters, K and Ω. As previously, a joint estimation of $2\sigma^2$, K and Ω is performed since $2\sigma^2$ is unknown. The likelihood function for joint ML estimation of $2\sigma^2$ and the Ricean distribution parameters can be derived by using integration as

$$f(\mathbf{y}_l|\sigma,\alpha,P) = \int_0^\infty \int_{-\pi}^{\pi} f(\mathbf{y}_l|\sigma,r_l,\theta_l) f_{R,\Theta}(r_l,\theta_l|\alpha,P)\,dr_l\,d\theta_l \tag{3.31}$$

where $f_{R,\Theta}(r_l,\theta_l|\alpha,P)$ is the joint PDF of the fading phase and the fading envelope in a Ricean channel. Solving (3.31), one has

$$\begin{aligned} f(\mathbf{y}_l|\sigma,\alpha,P) &= \left(\frac{1}{2\pi E\sigma^2}\right)^I \frac{2\sigma^2}{2\sigma^2+2\alpha^2} I_0\left(\frac{2z_lP}{2\sigma^2+2\alpha^2}\right) \\ &\quad \cdot \exp\left\{-\frac{1}{2E\sigma^2}\sum_{i=1}^{I}|y_{il}|^2+\frac{2\alpha^2 z_l^2}{2\sigma^2(2\sigma^2+2\alpha^2)}-\frac{P^2}{2\sigma^2+2\alpha^2}\right\}. \end{aligned} \quad (3.32)$$

If L independent symbols are used in the estimation, the likelihood function is then

$$\begin{aligned} f(\mathbf{y}|\sigma,\alpha,P) &= \left(\frac{1}{2\pi E\sigma^2}\right)^{LI} \left(\frac{2\sigma^2}{2\sigma^2+2\alpha^2}\right)^L \prod_{l=1}^{L} I_0\left(\frac{2z_lP}{2\sigma^2+2\alpha^2}\right) \\ &\quad \cdot \exp\left\{-\frac{1}{2E\sigma^2}\sum_{l=1}^{L}\sum_{i=1}^{I}|y_{il}|^2+\frac{2\alpha^2}{2\sigma^2(2\sigma^2+2\alpha^2)}\sum_{l=1}^{L} z_l^2-\frac{LP^2}{2\sigma^2+2\alpha^2}\right\}. \end{aligned}$$

(3.33)

Again, denote $\Omega' = P^2+2\alpha^2+2\sigma^2$ and $K' = \frac{P^2}{2\alpha^2+2\sigma^2}$. Rewriting (3.33) with Ω' and K', the ML estimators for $2\sigma^2$, Ω' and K' in a noisy Ricean fading channel can be derived as

$$2\hat{\sigma}^2 = \frac{1}{(I-1)LE}\sum_{l=1}^{L}\sum_{i=1}^{I}|y_{il}|^2 - \frac{1}{(I-1)L}\sum_{l=1}^{L} z_l^2, \quad (3.34)$$

$$\hat{\Omega}'_{Rice} = \frac{1}{L}\sum_{l=1}^{L} z_l^2 \quad (3.35)$$

and

$$\frac{1}{\hat{K}'_{ML}+1} - 1 - \frac{1}{L}\sum_{l=1}^{L} z_l'^2 + \frac{2\hat{K}'_{ML}+1}{L\sqrt{\hat{K}'_{ML}(\hat{K}'_{ML}+1)}}\sum_{l=1}^{L}\frac{I_1\left(2\sqrt{\hat{K}'_{ML}(\hat{K}'_{ML}+1)}z_l'\right)}{I_0\left(2\sqrt{\hat{K}'_{ML}(\hat{K}'_{ML}+1)}z_l'\right)} z_l' = 0 \quad (3.36)$$

where $z_l' = \frac{z_l}{\sqrt{\hat{\Omega}'_{Rice}}}$ and $I_1(\cdot)$ is the first-order modified Bessel function of the first kind [2, p. 374]. One sees that the ML estimates of $2\sigma^2$ and Ω', $2\hat{\sigma}^2$ and $\hat{\Omega}'_{Rice}$, can be calculated in explicit forms given by (3.34) and (3.35), respectively, while the ML estimate of K', $\hat{K}'_{ML}$, is implicitly determined by (3.36). According to the invariance principle of ML estimation [15], the ML estimates of K and Ω are, respectively,

$$\hat{K}_{ML} = \frac{\hat{K}'_{ML}\hat{\Omega}'_{Rice}}{\hat{\Omega}'_{Rice} - 2\hat{\sigma}^2(\hat{K}'_{ML}+1)} \quad (3.37a)$$

$$\hat{\Omega}_{Rice} = \hat{\Omega}'_{Rice} - 2\hat{\sigma^2}. \tag{3.37b}$$

Again, one sees that the ML estimators for Ω and K of the Ricean distribution derived when $2\sigma^2$ is unknown, $\hat{\Omega}_{Rice}$ and $\hat{K}_{ML}$, given in (3.37a) and (3.37b), respectively, have similar forms to those given in (3.8) derived when $2\sigma^2$ is known, except that the true value of $2\sigma^2$ in (3.8) is replaced by its ML estimate given in (3.34).

The mean and the variance of $2\hat{\sigma^2}$ in (3.34) are the same as those of the estimate given in (3.27). The CRLB for estimation of $2\sigma^2$ in a noisy Ricean channel can be derived as $\frac{1}{(I-1)L}\left(2\sigma^2\right)^2$. Therefore, the ML estimator for $2\sigma^2$ in a noisy Ricean fading channel is optimal in the sense of minimum variance, as was that in a noisy Rayleigh fading channel, as necessary since Rayleigh fading is a special case of Ricean fading. The mean and the variance of $\hat{\Omega}_{Rice}$ in (3.37b) can be derived as $E\{\hat{\Omega}_{Rice}\} = \Omega$ and $Var\{\hat{\Omega}_{Rice}\} = \frac{1}{L}(\frac{\Omega}{K+1} + 2\sigma^2)^2 + \frac{2K\Omega}{L(K+1)}(\frac{\Omega}{K+1} + 2\sigma^2) + \frac{1}{(I-1)L}\left(2\sigma^2\right)^2$. The CRLB for estimation of Ω in a noisy Ricean channel has to be calculated numerically. Next, we propose moment-based estimators for K in a noisy Ricean fading channel with unknown noise power. We use $\left\{z_l\right\}_{l=1}^{L}$ as noisy channel samples for moment-based estimation of K in this paper. One could use $\mathbf{y}_l$, samples of the received signal, as samples for moment-based estimation of K, as one did in the ML estimation of K. However, the choice of z_l is intuitively motivated as it was shown in Chapter 2 that z_l in (3.21) is a ML estimate of r_l.

To derive moment-based estimators for K when knowledge of the noise power in the channel is unavailable, we replace the true value of $2\sigma^2$ in the estimators given in the previous section with the ML estimate of $2\sigma^2$, $2\hat{\sigma^2}$, given in (3.34). Then, moment-based estimators for K using noisy channel samples when $2\sigma^2$ is unknown are given as

$$\hat{K}_{MB1} = \frac{\sqrt{2\hat{\mu}_2^2 - \hat{\mu}_4}}{\left(\hat{\mu}_2 - \sqrt{2\hat{\mu}_2^2 - \hat{\mu}_4}\right) - 2\hat{\sigma^2}} \tag{3.38}$$

and

$$\hat{K}_{MB2} = \frac{\hat{\mu}_2(3\hat{\mu}_2\hat{\mu}_1 - 2\hat{\mu}_3 + \hat{\mu}_2\sqrt{\hat{\mu}_1^2 - \hat{\mu}_{-1}\hat{\mu}_3 + \hat{\mu}_{-1}\hat{\mu}_1\hat{\mu}_2})}{\hat{\mu}_2(2\hat{\mu}_3 - 2\hat{\mu}_1\hat{\mu}_2) - 2\hat{\sigma^2}(\hat{\mu}_2\hat{\mu}_1 + \hat{\mu}_2\sqrt{\hat{\mu}_1^2 - \hat{\mu}_{-1}\hat{\mu}_3 + \hat{\mu}_{-1}\hat{\mu}_1\hat{\mu}_2})}. \tag{3.39}$$

Next, we develope estimators for the Nakagami-*m* distribution parameters.

3.2.4 Estimators for the Nakagami-*m* Distribution Parameters

Estimators for Nakagami-*m* distribution parameters, m and Ω, can be developed by using similar techniques to those used in the previous sections. We derive the likelihood function for joint ML estimation of $2\sigma^2$, m and Ω first. In a noisy Nakagami-*m* fading channel, this function can be obtained by solving the integration

$$f(\mathbf{y}_l|\sigma,m,\Omega) = \int_0^{\infty}\int_{-\pi}^{\pi} f(\mathbf{y}_l|\sigma,r_l,\theta_l) f_{R,\Theta}(r_l,\theta_l|m,\Omega)\,dr_l\,d\theta_l. \tag{3.40}$$

One has from (3.40)

$$f(\mathbf{y}_l|\sigma,m,\Omega) = \left(\frac{1}{2\pi E\sigma^2}\right)^{I}\left(\frac{2\sigma^2}{2\sigma^2+\frac{\Omega}{m}}\right)^{m} e^{-\frac{\sum_{i=1}^{I}|y_{il}|^2}{2E\sigma^2}}\cdot {}_1F_1\left(m,1;\frac{z_l^2}{2\sigma^2}-\frac{z_l^2}{2\sigma^2+\frac{\Omega}{m}}\right) \tag{3.41}$$

where ${}_1F_1(\cdot,\cdot;\cdot)$ is the confluent hypergeometric function [2, p. 504]. When using L independent symbols, the likelihood function for joint ML estimation is

$$\begin{aligned} f(\mathbf{y}|\sigma,m,\Omega) &= \left(\frac{1}{2\pi E\sigma^2}\right)^{LI}\left(\frac{2\sigma^2}{2\sigma^2+\frac{\Omega}{m}}\right)^{mL}\exp\left\{-\frac{1}{2E\sigma^2}\sum_{l=1}^{L}\sum_{i=1}^{I}|y_{il}|^2\right\} \\ &\quad \cdot\prod_{l=1}^{L} {}_1F_1\left(m,1;\frac{z_l^2}{2\sigma^2}-\frac{z_l^2}{2\sigma^2+\frac{\Omega}{m}}\right). \end{aligned} \tag{3.42}$$

The ML estimators for $2\sigma^2$, m and Ω might be found by using the log-likelihood function given in (3.42). Unfortunately, the difficulty in calculating the derivative of the log-likelihood function with respect to m makes it untractable. Therefore, moment-based estimators for m and Ω are proposed. Again, our moment-based estimation of Nakagami-*m* distribution parameters is based on samples $\{z_l\}_{l=1}^{L}$. The n-th order moment of the noisy Nakagami-*m* channel sample z_l is given in (3.15). Using the second order moment of z_l, one moment-based estimator for Ω, $\hat{\Omega}_{Naka}$, is

$$\hat{\Omega}_{Naka} = \hat{\Omega}'_{Naka} - 2\hat{\sigma}^2, \tag{3.43}$$

where $\hat{\Omega}'_{Naka} = \frac{1}{L}\sum_{l=1}^{L} z_l^2$ is the noiseless-based ML estimator for Ω mentioned in [75] and $2\sigma^2$ is the estimate of $2\sigma^2$ given by (3.27) and (3.34). This choice is motivated by the fact that the ML estimators for $2\sigma^2$ are the same for the Rayleigh and Ricean fading cases. Note that the moment-based estimator for Ω of the Nakagami-*m* distribution has a similar form to those ML estimators of

Ω for the Rayleigh and Ricean distributions. The mean of $\hat{\Omega}_{Naka}$ is $E\{\hat{\Omega}_{Naka}\} = \Omega$. The variance of $\hat{\Omega}_{Naka}$ is $Var\{\hat{\Omega}_{Naka}\} = \frac{1}{L}\left(8\sigma^2\Omega + 8\sigma^4 + \frac{\Omega^2}{m}\right) + \frac{1}{(I-1)L}\left(2\sigma^2\right)^2$. The CRLB for estimation of Ω in a noisy Nakagami-*m* fading channel has to be calculated numerically.

Using the moment-based estimator for Ω given in (3.43) and the estimator for $2\sigma^2$ in (3.27) and (3.34), moment-based estimators for *m* of the Nakagami-*m* distribution can also be developed. From [2, eqn. (15.2.11)], one has $(c-b)F(a,b-1;c;x) + (2b-c-bx+ax)F(a,b;c;x) + b(x-1)F(a,b+1;c;x) = 0$. Using this identitity, two equations can be obtained as

$$(\mu_2 - 2\sigma^2)\Omega + (6\sigma^2\mu_2 - 4\sigma^4 - \mu_4)m + \mu_2\Omega m = 0 \tag{3.44a}$$

$$(\mu_1 - \sigma^2\mu_{-1})\Omega + (8\sigma^2\mu_1 - 2\mu_3 - 2\sigma^4\mu_{-1})m + 2\mu_1 m\Omega = 0. \tag{3.44b}$$

Solving these two equations for *m* and using the moment-based estimator for Ω in (3.43) and the estimator for $2\sigma^2$ in (3.27) and (3.34), two moment-based estimators for *m*, $\hat{m}_{MB1}$ and $\hat{m}_{MB2}$, are derived as

$$\hat{m}_{MB1} = \frac{(\hat{\mu}_2 - 2\hat{\sigma}^2)^2}{\hat{\mu}_4 - \hat{\mu}_2^2 - 4\hat{\sigma}^2\hat{\mu}_2 + 4\hat{\sigma}^4} \tag{3.45}$$

and

$$\hat{m}_{MB2} = \frac{(\hat{\mu}_1 - \hat{\sigma}^2\hat{\mu}_{-1})(\hat{\mu}_2 - 2\hat{\sigma}^2)}{2\hat{\mu}_3 - 2\hat{\mu}_1\hat{\mu}_2 + 2\hat{\sigma}^4\hat{\mu}_{-1} - 4\hat{\sigma}^2\hat{\mu}_1}, \tag{3.46}$$

where $\hat{\mu}_n$ is defined as before. Note that (3.46) uses the first, the second, the third and the inverse first order moments of the noisy samples, while (3.45) uses the second and the fourth order moments. Therefore, it is expected that (3.46) will outperform (3.45), as the former uses a lower order of moment which is more robust to noise outlyers. The noiseless-sample-based estimators derived in the literature are [74]

$$\hat{m}_s = \frac{\hat{\mu}_2^2}{\hat{\mu}_4 - \hat{\mu}_2^2} \tag{3.47}$$

and

$$\hat{m}_t = \frac{\hat{\mu}_1\hat{\mu}_2}{2\hat{\mu}_3 - 2\hat{\mu}_1\hat{\mu}_2}. \tag{3.48}$$

We will compare the performances of $\hat{m}_{MB1}$ and $\hat{m}_{MB2}$ with those of $\hat{m}_s$ and $\hat{m}_t$ later.

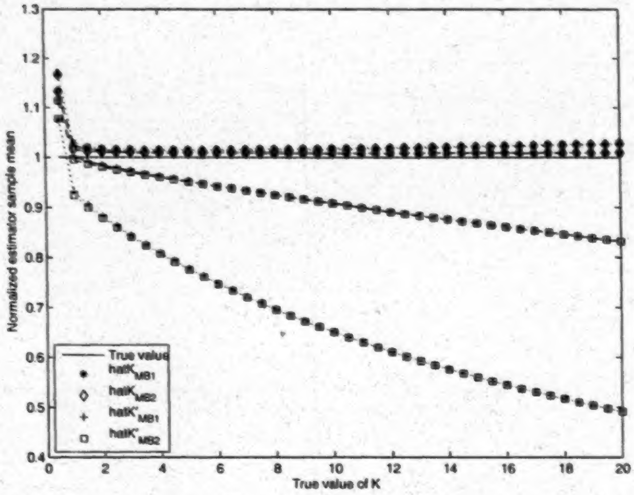

Figure 3.12. Comparison of the normalized sample means of the estimators $\hat{K}_{MB1}$ in (3.38), $\hat{K}_{MB2}$ in (3.39), $\hat{K'}_{MB1}$ in (3.10) and $\hat{K'}_{MB2}$ in (3.12) with the true value for $L = 500$, TSNR = 7 dB (dashed line), $L = 500$, TSNR = 0 dB (dotted line) in a noisy Ricean fading channel.

3.2.5 Numerical Results

In this part, numerical results are presented to show the performances of the estimators derived. The value of I is set equal to 2, which corresponds to the case of two-dimensional signaling. Consider $\sigma^2 = 0.1$ (TSNR = 7 dB) and $\sigma^2 = 0.5$ (TSNR = 0 dB).

Figs. 3.12 to 3.15 show the performances of $\hat{K}_{MB1}$ in (3.38), $\hat{K}_{MB2}$ in (3.39), $\hat{K'}_{MB1}$ in (3.10) and $\hat{K'}_{MB2}$ in (3.12) for the Ricean distribution parameter K, where negative estimates have been set equal to zero. The estimators $\hat{K}_{MB1}$ and $\hat{K}_{MB2}$ are noisy-sample-based while the estimators $\hat{K'}_{MB1}$ and $\hat{K'}_{MB2}$ are noiseless-sample-based. Figs. 3.12 and 3.13 show the normalized sample means of the estimators, where the sample means are normalized with respect to the true value. One sees that, with the same sample size, the biases of the estimators decrease significantly as the TSNR increases, while with the same TSNR, the biases of the estimators decrease little even when the sample size increases. Therefore, the TSNR has a greater influence on the sample means than the sample size.

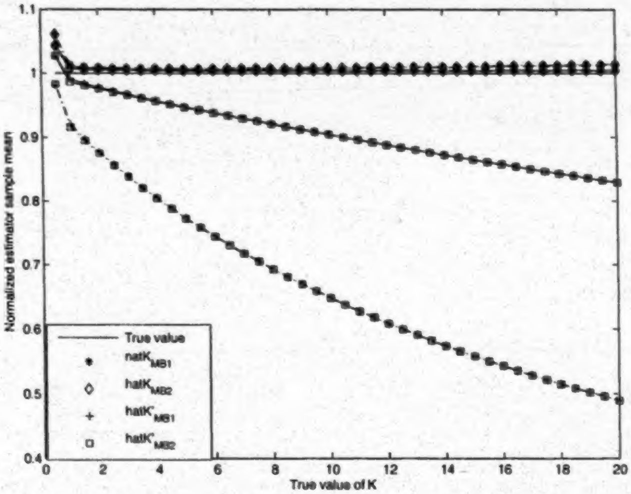

Figure 3.13. Comparison of the normalized sample means of the estimators $\hat{K}_{MB1}$ in (3.38), $\hat{K}_{MB2}$ in (3.39), $\hat{K}'_{MB1}$ in (3.10) and $\hat{K}'_{MB2}$ in (3.12) with the true value for $L = 1000$, TSNR = 7 dB (solid line), $L = 1000$, TSNR = 0 dB (dashdotted line) in a noisy Ricean fading channel.

In the case when both the sample size and the TSNR are fixed, $\hat{K}_{MB1}$ and $\hat{K}_{MB2}$ have much smaller biases than $\hat{K}'_{MB1}$ and $\hat{K}'_{MB2}$. Moreover, the biases of $\hat{K}'_{MB1}$ and $\hat{K}'_{MB2}$ increase dramatically as the true value of K increases while the biases of $\hat{K}_{MB1}$ and $\hat{K}_{MB2}$ changes little. For example, when $L = 500$ and TSNR = 0 dB, $\hat{K}_{MB1}$ and $\hat{K}_{MB2}$ have a bias which is about +1% of the true value at $K = 2$ and +2% at $K = 20$. However, $\hat{K}'_{MB1}$ and $\hat{K}'_{MB2}$ have a bias which is about -12% of the true value at $K = 2$ and -51% of the true value at $K = 20$. Figs. 3.14 and 3.15 show the root mean squared error (RMSE) of the estimators. The CRLBs for $K = 0$ go to infinity, and they are not shown in the figures. The RMSE decreases significantly as the TSNR increases and the sample size remains the same. When the sample size increases and the TSNR remains the same, the RMSE decreases little. Therefore, the TSNR also has a greater influence on the RMSE than the sample size. If both the sample size and the TSNR are fixed, one sees that $\hat{K}_{MB1}$ and $\hat{K}_{MB2}$ have much smaller RMSEs as well as much smaller deviations from the CRLBs than $\hat{K}'_{MB1}$ and $\hat{K}'_{MB2}$. For

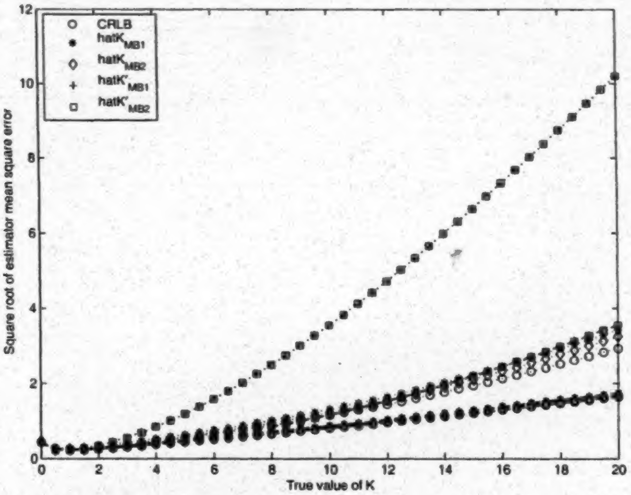

Figure 3.14. Comparison of the RMSEs of the estimators $\hat{K}_{MB1}$ in (3.38), $\hat{K}_{MB2}$ in (3.39), $\hat{K}'_{MB1}$ in (3.10) and $\hat{K}'_{MB2}$ in (3.12) with the CRLBs for $L = 500$, TSNR = 7 dB (dashed line), $L = 500$, TSNR = 0 dB (dotted line) in a noisy Ricean fading channel.

example, when $L = 500$ and TSNR = 0 dB, $\hat{K}_{MB1}$ and $\hat{K}_{MB2}$ have approximately the same RMSE, which is about 0.7 at $K = 6$ and about 3.3 at $K = 20$. The deviation from the CRLB is negligible at $K = 6$ and about 0.4 at $K = 20$. However, $\hat{K}'_{MB1}$ and $\hat{K}'_{MB2}$ have a RMSE of about 1.6 at $K = 6$ and about 10.1 at $K = 20$. The deviation from the CRLB is about 0.9 at $K = 6$ and about 7.1 at $K = 20$. Therefore, the noisy-sample-based estimators, $\hat{K}_{MB1}$ and $\hat{K}_{MB2}$, perform much better than the noiseless-sample-based estimators, $\hat{K}'_{MB1}$ and $\hat{K}'_{MB2}$. Comparing the two noisy-sample-based estimators, $\hat{K}_{MB1}$ and $\hat{K}_{MB2}$, one sees that they have similar sample mean performance while the latter has a smaller RMSE when the sample size and the TSNR are the same. Therefore, $\hat{K}_{MB2}$ outperforms $\hat{K}_{MB1}$, as expected.

Figs. 3.16 and 3.19 show the performances of $\hat{m}_{MB1}$ in (3.45), $\hat{m}_{MB2}$ in (3.46), $\hat{m}_s$ in (3.47) and $\hat{m}_t$ in (3.48) for the Nakagami m parameter, where estimates less than 0.5 have been set equal to 0.5. The first two are noisy-sample-based while the last two are noiseless-sample-based. For

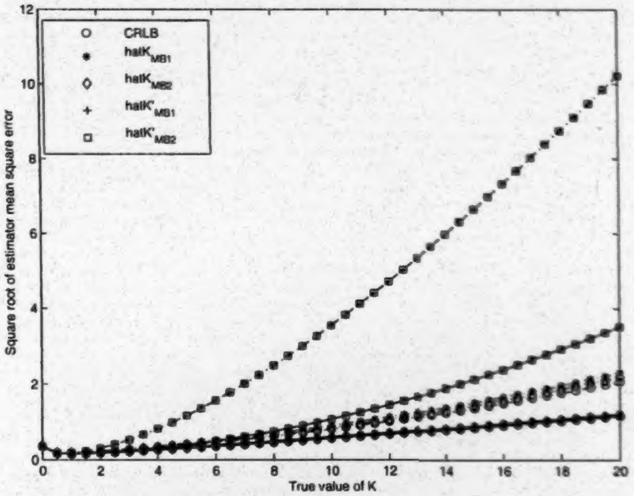

Figure 3.15. Comparison of the RMSEs of the estimators $\hat{K}_{MB1}$ in (3.38), $\hat{K}_{MB2}$ in (3.39), $\hat{K}'_{MB1}$ in (3.10) and $\hat{K}'_{MB2}$ in (3.12) with the CRLBs for $L = 1000$, TSNR = 7 dB (solid line), $L = 1000$, TSNR = 0 dB (dashdotted line) in a noisy Ricean fading channel.

comparison, performances of the noisy-sample-based m parameter estimators derived in [76, eqn. (28)] and [79, eqn. (14)], denoted as $\hat{m}_{old1}$ and $\hat{m}_{old2}$, respectively, are also shown in Figs. 3.17 and 3.19 for $L = 10,000$ and TSNR=7 dB. The derivation of the CRLB for m is untractable in a noisy Nakagami-m channel. The CRLB in the noiseless case, which is a lower bound of that in the noisy case, is used instead in Figs. 3.18 and 3.19. Figs. 3.16 and 3.17 show the normalized sample means of the estimators. Similar observations to those made for the K parameter estimators can be made. Again, the biases decrease as the TSNR and/or the sample size increases, and the TSNR also has a greater influence on the sample means than the sample size. When the sample size and the TSNR are the same, the noisy-sample-based estimators have much smaller biases than the noiseless-sample-based estimators and the biases of the noisy-sample-based estimator also increases much slower than the noiseless-sample-based ones when m increases. As an example, when $L = 500$ and TSNR = 0 dB, $\hat{m}_{MB1}$ and $\hat{m}_{MB2}$ have approximately the same positive bias of about +2% at $m = 6$ and about

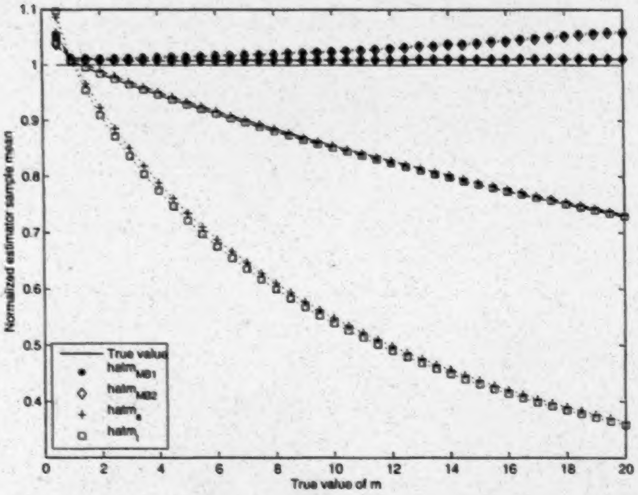

Figure 3.16. Comparison of the normalized sample means of $\hat{m}_{MB1}$ in (3.45), $\hat{m}_{MB2}$ in (3.46), $\hat{m}_s$ in (3.47) and $\hat{m}_t$ in (3.48) for $L = 500$, TSNR = 7 dB (dashed line), $L = 500$, TSNR = 0 dB (dotted line) in a noisy Nakagami-m fading channel.

+5% at $m = 20$, while $\hat{m}_s$ and $\hat{m}_t$ have a negative bias of about -33% of the true value at $m = 6$ and -63% at $m = 20$. Figs. 3.18 and 3.19 show the RMSEs of the estimators. Again, the RMSE decreases as the TSNR and/or the sample size increases, and the TSNR has a greater influence. With the same TSNR and sample size, $\hat{m}_{MB1}$ and $\hat{m}_{MB2}$ have much smaller RMSEs than $\hat{m}_s$ and $\hat{m}_t$. This can be seen when $L = 500$ and TSNR = 0 dB. The noisy-sample-based estimators $\hat{m}_{MB1}$ and $\hat{m}_{MB2}$ have RMSE of about 0.8 at $m = 6$ and about 5.5 at $m = 20$ while the noiseless-sample-based estimators $\hat{m}_s$ and $\hat{m}_t$ have RMSE of about 2.1 at $m = 6$ and about 12.0 at $m = 20$. Therefore, the noisy-sample-based estimators, $\hat{m}_{MB1}$ and $\hat{m}_{MB2}$, outperform the noiseless-sample-based estimators, $\hat{m}_s$ and $\hat{m}_t$. Also, comparing the two noisy-sample-based estimators, $\hat{m}_{MB1}$ and $\hat{m}_{MB2}$, one sees that they have nearly the same sample mean performance while the latter has a smaller RMSE. Therefore, $\hat{m}_{MB2}$ outperforms $\hat{m}_{MB1}$, as expected. One can also see from Figs. 3.17 and 3.19 that our noisy-sample-based estimators have smaller biases as well as smaller RMSEs at $L = 1000$ and TSNR = 0 dB than

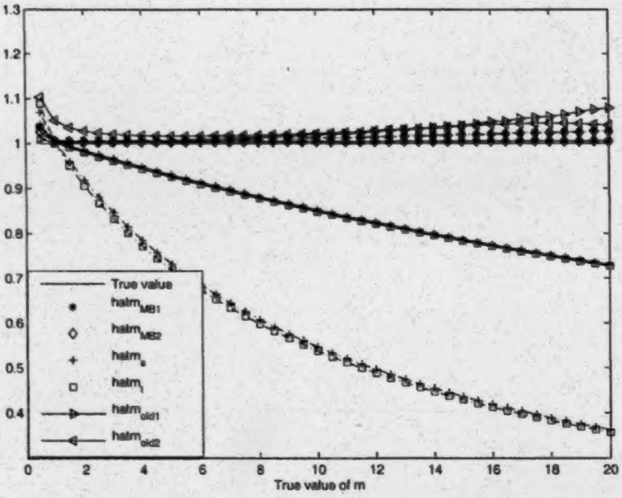

Figure 3.17. Comparison of the normalized sample means of $\hat{m}_{MB1}$ in (3.45), $\hat{m}_{MB2}$ in (3.46), $\hat{m}_s$ in (3.47) and $\hat{m}_t$ in (3.48) for $L = 1000$, TSNR = 7 dB (solid line), $L = 1000$, TSNR = 0 dB (dashdotted line) in a noisy Nakagami-*m* fading channel. $\hat{m}_{old1}$ and $\hat{m}_{old2}$ use $L = 10,000$ and TSNR = 7 dB.

the noisy-sample-based estimators derived in references [76] and [79] at $L = 10,000$ and TSNR = 7 dB, which implies that they are at least ten times better than the noisy-sample-based estimators in [76] and [79] even when stronger noise is assumed. Therefore, our noisy-sample-based estimators for the *m* parameter perform better than not only the noiseless-sample-based estimators but also the noisy-sample-based estimators for *m* parameter derived in the literature when both are used in a noisy Nakagami-*m* fading channel.

3.3 Channel Model Parameter Estimation Using Phase Samples

In the previous two sections, estimators for fading distribution parameters with or without knowledge of the noise power have been studied. All these estimators, including those derived in the literature, use samples of the fading envelope (the absolute value of the complex fading gain). However, in some cases, the fading phase (the angle of the complex fading gain) contains information

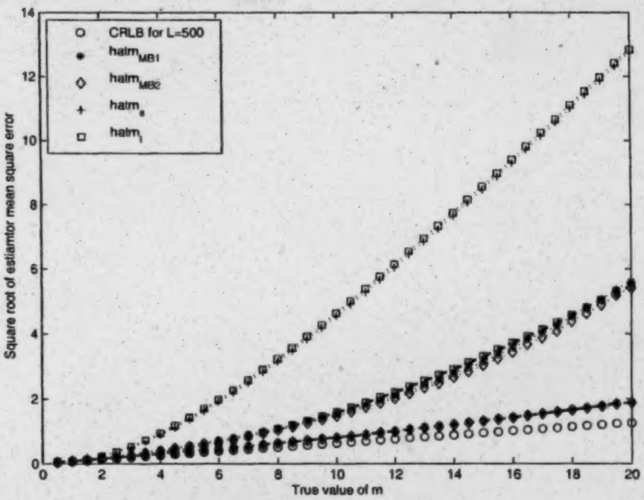

Figure 3.18. Comparison of the RMSEs of $\hat{m}_{MB1}$ in (3.45), $\hat{m}_{MB2}$ in (3.46), $\hat{m}_s$ in (3.47) and $\hat{m}_t$ in (3.48) with the CRLBs for $L = 500$, TSNR = 7 dB (dashed line), $L = 500$, TSNR = 0 dB (dotted line) in a noisy Nakagami-*m* fading channel.

on the fading distribution parameters as well. It is of great interest to explore the use of the fading phase in the estimation of the fading distribution parameters. In this work, we propose ML and approximate ML estimators for the Ricean K parameter by using fading phase samples.

3.3.1 System Model

It is well known that the complex fading gain in a Ricean fading channel can be modeled as a Gaussian random variable according to the central limit theorem [54]. The joint probability density function (PDF) of the fading envelope and the fading phase is given by [83]

$$f_{R,\Theta}(r,\theta) = \frac{r}{2\pi\alpha^2}\exp\left\{-\frac{r^2+P^2-2rP\cos(\theta-\theta_0)}{2\alpha^2}\right\} \tag{3.49}$$

where $r > 0$ is the fading envelope, $-\pi < \theta < \pi$ is the fading phase, P^2 is the mean power of the LOS component, $2\alpha^2$ is the mean power of the scattering components, and θ_0 is the angle of arrival of the LOS component with $-\pi < \theta_0 < \pi$. Define $K = \frac{P^2}{2\alpha^2}$ and $\Omega = P^2 + 2\alpha^2$ as the Ricean K factor

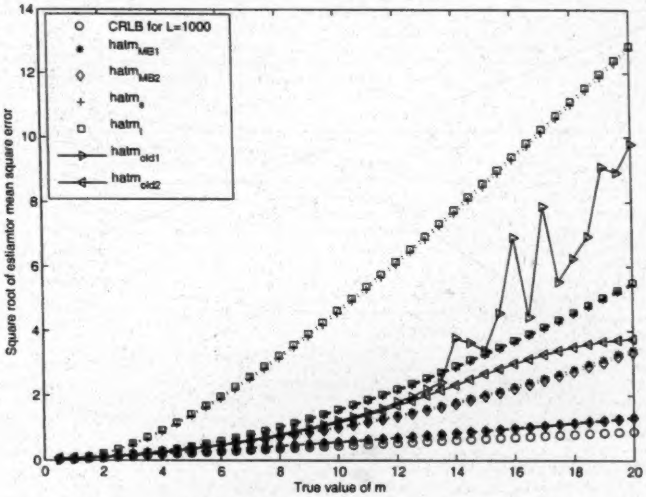

Figure 3.19. Comparison of the RMSEs of $\hat{m}_{MB1}$ in (3.45), $\hat{m}_{MB2}$ in (3.46), $\hat{m}_s$ in (3.47) and $\hat{m}_t$ in (3.48) with the CRLBs for $L = 1000$, TSNR = 7 dB (solid line), $L = 1000$, TSNR = 0 dB (dashdotted line) in a noisy Nakagami-m fading channel. $\hat{m}_{old1}$ and $\hat{m}_{old2}$ use $L = 10,000$ and TSNR = 7 dB.

and the total mean power of the fading, respectively, as before. By integrating over θ in (3.49), one obtains the fading envelope Ricean PDF as (1.3). By integrating over r in (3.49), one obtains the fading phase PDF

$$f_\Theta(\theta) \;=\; \frac{e^{-K}}{2\pi} + \frac{\sqrt{K}\cos(\theta-\theta_0)}{2\sqrt{\pi}} e^{-K\sin^2(\theta-\theta_0)} \cdot \operatorname{erfc}(-\sqrt{K}\cos(\theta-\theta_0)) \tag{3.50}$$

where $\operatorname{erfc}(\cdot)$ is the complementary error function.

3.3.2 K Estimator Using Phase Samples Only

When K is large, one has $e^{-K} \approx 0$. From (3.50), one derives an approximation to the log-likelihood function using samples of the fading phase only as

$$\begin{aligned} \ln f(\theta|K,\theta_0) \;\approx\; & -L\ln(2\sqrt{\pi}) + \sum_{l=1}^{L} \ln\cos(\theta_l-\theta_0) - K\sum_{l=1}^{L}\sin^2(\theta_l-\theta_0) + \frac{L}{2}\ln K \\ & + \sum_{l=1}^{L} \ln \operatorname{erfc}(-\sqrt{K}\cos(\theta_l-\theta_0)) \end{aligned} \tag{3.51}$$

where $\theta = [\theta_1 \quad \theta_2 \quad \cdots \quad \theta_L]$ is the sample vector of the fading phase and the term $\frac{e^{-K}}{2\pi}$ in (3.50) has been ignored. It can also be shown that $\mathrm{erfc}(-\sqrt{K}\cos(\theta_l - \theta_0))$ depends only on the sign of $\cos(\theta_l - \theta_0)$ and that $E\{\theta\} \approx \theta_0$, for large values of K. As a result, an approximate ML-based estimator for K can be obtained from (3.51) as

$$\hat{K}_{AML} = \frac{L}{2\sum_{l=1}^{L} \sin^2(\theta_l - \hat{\theta}_0)} \tag{3.52}$$

where $\hat{\theta}_0 = \frac{1}{L}\sum_{l=1}^{L} \theta_l$. One sees that, unlike the ML estimator for K in [70] that uses samples of r_l only, the estimator in (3.52) has a very simple structure.

3.3.3 K Estimator Using Both Phase Samples and Envelope Samples

From (3.49), the log-likelihood function for estimation of K using both fading envelope samples and fading phase samples can be derived as

$$\ln f(\mathbf{r}, \theta | P, \alpha, \theta_0) = -\frac{\sum_{l=1}^{L} r_l^2 - 2P\sum_{l=1}^{L} r_l \cos(\theta_l - \theta_0)}{2\alpha^2} + \sum_{l=1}^{L} \ln r_l - L\ln(2\pi\alpha^2) - \frac{LP^2}{2\alpha^2} \tag{3.53}$$

where $\mathbf{r} = [r_1 \quad r_2 \quad \cdots \quad r_L]$ is the sample vector of the fading envelope. By differentiating (3.53) with respect to θ_0, P, $2\alpha^2$, setting the derivatives equal to zero, and solving the resulting equations, one has the ML estimators for θ_0, P and $2\alpha^2$, $\hat{\theta}_0$, $\hat{P}$ and $2\hat{\alpha}^2$, as

$$\hat{\theta}_0 = \arctan\frac{\sum_{l=1}^{L} r_l \sin\theta_l}{\sum_{l=1}^{L} r_l \cos\theta_l} \tag{3.54}$$

$$\hat{P} = \frac{1}{L}\sum_{l=1}^{L} r_l \cos(\theta_l - \hat{\theta}_0) \tag{3.55}$$

$$2\hat{\alpha}^2 = \frac{1}{L}\sum_{l=1}^{L} r_l^2 - \hat{P}^2. \tag{3.56}$$

According to the invariance principle of ML estimation [15], the ML estimator for the Ricean K factor, $\hat{K}_{ML}$, is

$$\hat{K}_{ML} = \frac{\hat{P}^2}{2\hat{\alpha}^2} \tag{3.57}$$

where $\hat{P}$ is given by (3.55) and $2\hat{\alpha}^2$ is given by (3.56). One sees that (3.57) uses samples of the fading envelope, r_l, as well as samples of the fading phase, θ_l. The ML estimator in (3.57) also has a simple structure which is implementable.

Finally, the best known realizable noiseless-based estimator for K using samples of the fading envelope only, $\hat{K}_{Env}$, is derived in the literature by solving the equation [70]

$$\frac{\hat{\mu}_1}{\sqrt{\hat{\mu}_2}} = \frac{\sqrt{\pi}e^{-K/2}}{2\sqrt{K+1}}\left[(1+K)I_0(K/2)+KI_1(K/2)\right] \tag{3.58}$$

for K, where $\hat{\mu}_1 = \frac{1}{L}\sum_{l=1}^{L} r_l$, $\hat{\mu}_2 = \frac{1}{L}\sum_{l=1}^{L} r_l^2$, and $I_1(\cdot)$ is the first-order modified Bessel function of the first kind. One sees that, unlike the new estimators in (3.52) and (3.57), this estimator doesn't have an explicit form. In practice, it has to be implemented by using a look-up table, which costs extra memory as well as searching time. A closed-form estimator for K based on an approximation to (3.58) was derived in [72].

3.3.4 Numerical Results

In this section, we simulate the performances of $\hat{K}_{ML}$ and $\hat{K}_{AML}$ and compare them with that of $\hat{K}_{Env}$ in a Ricean fading channel. The Ricean channel is simulated by generating independent complex Gaussian random variables with, in general, nonzero mean. Thus, the scattering distribution of the incoming waves doesn't contribute any error to the estimate. Without loss of generality, we fix the angle of arrival of the LOS component to 1.0 and the total mean power of the fading to 1.0. The value of K varies from 0.0 to 10.0 with increments of 0.5. Practical values of K are reported to be less than 10.0 in mobile communication systems [71]. The look-up table in $\hat{K}_{Env}$ is constructed for K from 0.0 to 15.0 with a step size of 0.05, resulting in 300 possible values for search.

Figs. 3.20 and 3.21 show the biases and the RMSEs of $\hat{K}_{ML}$, $\hat{K}_{AML}$ and $\hat{K}_{Env}$ in a Ricean fading channel with $L = 500$. One sees that the bias of $\hat{K}_{ML}$ is between 0.00 and 0.05, and the RMSE of $\hat{K}_{ML}$ is between 0.00 and 0.50, when the true value of K varies from 0.0 to 10.0. The bias and the RMSE of $\hat{K}_{ML}$ are very small. As an example, at $K = 10.0$, the bias of $\hat{K}_{ML}$ is only 0.5% of the true value and the RMSE of $\hat{K}_{ML}$ is only 5% of the true value. Therefore, the estimator $\hat{K}_{ML}$ performs

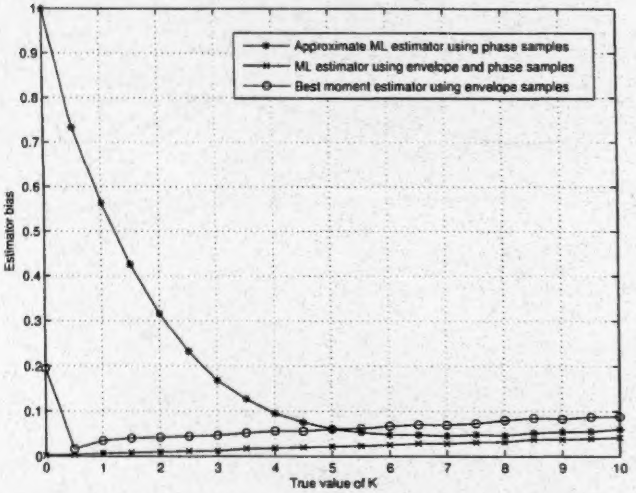

Figure 3.20. The biases of $\hat{K}_{ML}$, $\hat{K}_{AML}$ and $\hat{K}_{Env}$ for a sample size of $L = 500$ in a Ricean fading channel.

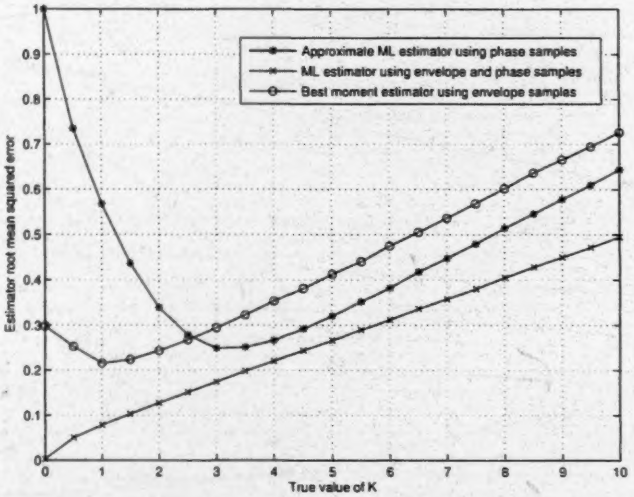

Figure 3.21. The RMSEs of $\hat{K}_{ML}$, $\hat{K}_{AML}$ and $\hat{K}_{Env}$ for a sample size of $L = 500$ in a Ricean fading channel.

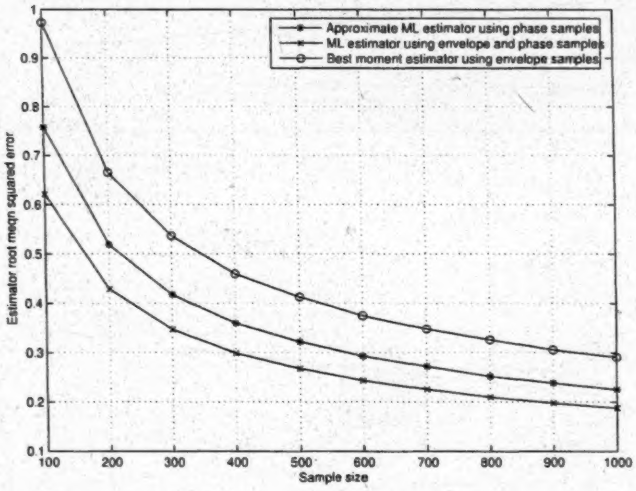

Figure 3.22. The RMSEs of $\hat{K}_{ML}$, $\hat{K}_{AML}$ and $\hat{K}_{Env}$ versus the sample size with $K = 5$ in a Ricean fading channel.

well in a Ricean fading channel. Also, one sees from Figs. 3.20 and 3.21 that the estimator $\hat{K}_{AML}$ has good performance when the true value of K is larger than 3.0; the bias of $\hat{K}_{AML}$ is between 0.17 and 0.06 and the RMSE of $\hat{K}_{AML}$ is between 0.25 and 0.64 for $3.0 \leq K \leq 10.0$. However, when $0.0 \leq K < 3.0$, $\hat{K}_{AML}$ performs poorly, as the bias and the RMSE of $\hat{K}_{AML}$ are fairly large, and they increase when the true value of K decreases. This is caused by the large errors in the approximations used to derive (3.52) at small values of K. Comparing $\hat{K}_{ML}$ with $\hat{K}_{Env}$, one sees that $\hat{K}_{ML}$ always has a smaller bias as well as a smaller RMSE than $\hat{K}_{Env}$, for all the values of K considered. Thus, $\hat{K}_{ML}$ outperforms $\hat{K}_{Env}$. Comparing $\hat{K}_{AML}$ with $\hat{K}_{Env}$, one sees that $\hat{K}_{AML}$ outperforms $\hat{K}_{Env}$ when $5.0 \leq K \leq 10.0$. When $0.0 \leq K < 2.5$, $\hat{K}_{AML}$ underperforms $\hat{K}_{Env}$. Otherwise, their performances are comparable. Fig. 3.22 shows how the RMSEs of the estimators vary with the sample size at $K = 5$. One can determine how much the estimator performance improves as the the sample size increases from Fig. 3.22. For example, $\hat{K}_{ML}$ has a RMSE of 0.62 when $L = 100$, while it has a RMSE of 0.19 when $L = 1000$. Next, we investigate the problem of signal-to-noise ratio estimation.

Chapter 4

SNR Estimation in Wireless Communication Channel

As we mentioned in Chapter 1, signal-to-noise ratio (SNR) is an important channel parameter that is widely used in wireless communication. Two measures of SNR can be defined in a static additive white Gaussian noise (AWGN) channel. One is the SNR defined as the ratio of the signal power to the noise power. The other is the SNR defined as the ratio of the signal amplitude to the noise power. In a fading channel, a third SNR, the average signal-to-noise-ratio (ASNR), is also defined as the ratio of the average signal power to the noise power. In contrast, the first two SNR measures are refered to as instantaneous SNRs. In the following, until otherwise specified, the term of SNR refers to the instantaneous SNRs in a fading channel. Many techniques and components need SNR and ASNR information in their implementation. For example, knowledge of SNR and ASNR is required in rate adaptation [13], power control [11], [84], optimum diversity combining [85], macro-diversity [86], turbo decoding [87], and maximal ratio combining (MRC) [88]. These applications give impetus to an investigation of their estimation techniques. Previous works include the following. In [19]- [21], maximum likelihood (ML) SNR estimators were derived and shown to be asymptotically optimal in the sense of minimum variance. These results are only applicable for pulse-code modulated signals in a real AWGN channel. Pauluzzi and Beaulieu extended these results to M-ary phase-shift keying (MPSK) signals in complex AWGN channels [22]. They also summarized and compared their predecessors' work in [22]. In another work [23], based on observations from the link quality estimator used in [24], Beaulieu *et al.* proposed four different SNR estimators for quaternary phase-shift keying (QPSK) signals. These estimators have fairly large biases when the true value of SNR is small. To improve the estimator bias performance at small SNRs, Li *et al.* developed another estimator in [25]. In related work [26], [27], the Cramér-Rao lower bound (CRLB) for non-data-aided SNR estimation of phase-shift keying signals in an AWGN channel was obtained. A new non-data-aided SNR estimator based on an iterative algorithm was also derived in [27]. All these estimators are limited to a static AWGN channel. They are only valid for a sampled system with bandlimited AWGN. The performances of most estimators were examined by simulation. Estimation of ASNR in a slowly fading channel has been studied by many researchers as well. In [81], [89], [90], estimation of the average signal power has been studied. These estimators can be adapted to estimate ASNR assuming that the noise power is known. In [91] and [92], moment-based estimators for ASNR in a Rayleigh fading channel and a Nakagami-*m* fading channel, respectively, were derived. There is no result on estimation of ASNR in a Ricean fading channel without knowledge of the noise power though. In this chapter, we first derive ML estimators for SNR in a static AWGN channel as well as a slowly fading channel, considering both sampled system with bandlimited AWGN and continuous system. The performances of the estimators are examined by analysis. We then design approximate ML estimators for SNR using both pilot and data symbols. We also derive a ML estimator for ASNR in a Ricean fading channel without knowledge of the noise power and perform a joint estimation of ASNR and the Ricean K parameter using noisy correlated channel samples. Finally, we design moment-based estimators for the signal-to-interference-plus-noise ratio (SINR) when interferences occur and quadrature amplitude modulation (QAM) is used.

4.1 ML Decision-Based Estimation of SNR

In this section, we derive ML decision-based estimators for SNR in static AWGN and slowly fading memoryless channels. For convenience, we denote $\rho_1 = \frac{A}{N}$ as SNRA and $\rho_2 = \frac{P_s}{N}$ as SNRB, where A is the channel gain, N is the noise power, and P_s is the signal power. The ML estimators for A, N

and P_s have been derived in Chapter 2. Using these results, together with the invariance principle, the ML estimators for ρ_1 and ρ_2 can be derived.

4.1.1 Static AWGN Channel

We consider first a static AWGN channel. According to the invariance principle, one has

$$\hat{\rho}_1 = \frac{\hat{A}}{\hat{N}} \tag{4.1}$$

and

$$\hat{\rho}_2 = c\frac{\hat{A}^2}{\hat{N}} \tag{4.2}$$

where $\hat{A}$ and $\hat{N}$ are the ML estimates of A and N in a static AWGN channel, respectively, $c = \frac{E_{sd}^{(\hat{k})}}{I}$ for a sampled system and $c = \frac{E_{sc}^{(\hat{k})}}{T}$ for a continuous time system, and $E_{sd}^{(\hat{k})}$, $E_{sc}^{(\hat{k})}$, I and T are defined in Section 2.1. From (2.5) and (2.6), one has

$$\hat{\rho}_1 = \frac{I\sum_{i=1}^{I} Re\{y_i s_i^{(\hat{k})*}\}}{E_{sd}^{(\hat{k})}\left(\sum_{i=1}^{I}|y_i|^2\right) - \left[\sum_{i=1}^{I} Re\{y_i s_i^{(\hat{k})*}\}\right]^2} \tag{4.3}$$

and

$$\hat{\rho}_2 = \frac{\left[\sum_{i=1}^{I} Re\{y_i s_i^{(\hat{k})*}\}\right]^2}{E_{sd}^{(\hat{k})}\left(\sum_{i=1}^{I}|y_i|^2\right) - \left[\sum_{i=1}^{I} Re\{y_i s_i^{(\hat{k})*}\}\right]^2} \tag{4.4}$$

as the sampled signal ML estimators for SNRA and SNRB, respectively. Also, from (2.8) and (2.10), one has

$$\hat{\rho}_1 = \frac{T\int_T Re\{y(t)s^{(\hat{k})*}(t)\}dt}{E_{sc}^{(\hat{k})}\left(\int_T |y(t)|^2 dt\right) - \left[\int_T Re\{y(t)s^{(\hat{k})*}(t)\}dt\right]^2} \tag{4.5}$$

and

$$\hat{\rho}_2 = \frac{\left[\int_T Re\{y(t)s^{(\hat{k})*}(t)\}dt\right]^2}{E_{sc}^{(\hat{k})}\left(\int_T |y(t)|^2\right) - \left[\int_T Re\{y(t)s^{(\hat{k})*}(t)\}dt\right]^2}. \tag{4.6}$$

as the contimuous signal ML estimators for SNRA and SNRB, respectively.

Fig. 4.1 shows a block diagram of the ML SNRA and SNRB estimators given by (4.5) and (4.6), respectively. One sees that the ML static AWGN channel SNR estimator for a digitally modulated signal is a decision-based structure that incorporates a digital data receiver. The bulk of the cost and

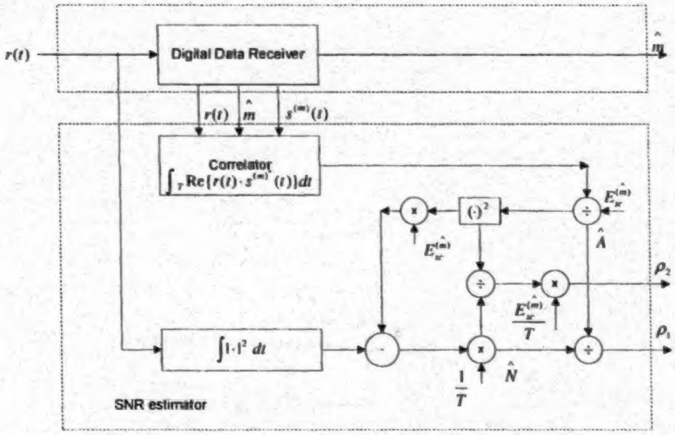
$r(t)$
Digital Data Receiver
$\hat{m}$
$r(t)$
$\hat{m}$
$s^{(m)}(t)$
Correlator
$\int_T \mathrm{Re}\{r(t)\cdot s^{(m)}(t)\}dt$
$(\cdot)^2$
$E_{sr}^{(\hat{m})}$
$E_{sr}^{(\hat{m})}$
$\hat{A}$
ρ_2
$\frac{E_{sr}^{(\hat{m})}}{T}$
ρ_1
$\int|\cdot|^2\,dt$
$\frac{1}{T}$
$\hat{N}$
SNR estimator

Figure 4.1. Structures of the ML SNR estimators for ρ_1 and ρ_2 in an AWGN channel.

complexity of the ML estimator resides in the digital receiver component, used to determine $\hat{k}$. The digital receiver is required anyway. Thus, ML SNR estimation is achievable with relatively minor complexity. Next, we analyze the performances of $\hat{\rho}_1$ and $\hat{\rho}_2$ in error-free operation.

For two independent random variables T_1 and T_2, the PDF of $T = \frac{T_1}{T_2}$ can be derived by solving the integral [34, eq. (6-60)]

$$p_T(t) = \int_0^{\infty} t_2 p_{T_1}(t_2 t) p_{T_2}(t_2) dt_2. \tag{4.7}$$

By using the results of (2.18) and (2.25), together with the identity [35, eq. (3.462.1)], one has the PDF of the sampled signal ML estimate of ρ_1 as

$$p_{\hat{\rho}_1}(x) = \frac{2I-1}{2\sqrt{2\pi x}} \left(\frac{I\sigma}{Nx}\right)^{I-\frac{1}{2}} e^{\frac{1}{4}\left(\frac{I\sigma}{Nx}-\frac{A}{\sigma}\right)^2 - \frac{A^2}{2\sigma^2}} D_{-(I+\frac{1}{2})}\left(\frac{I\sigma}{Nx} - \frac{A}{\sigma}\right) \tag{4.8}$$

where $\sigma^2 = \frac{N}{2E_{sd}^{(k)}}$ and $D_v(\cdot)$ is the parabolic cylinder function [2, p. 686]. The mean and the variance of $\hat{\rho}_1$ are $E\{\hat{\rho}_1\} = \frac{2I}{2I-3}\rho_1$ and $Var\{\hat{\rho}_1\} = \frac{8I^2}{(2I-3)^2(2I-5)}\rho_1^2 + \frac{4I^2}{(2I-3)(2I-5)}\frac{1}{2NE_{sd}^{(k)}}$. The Cramér-Rao lower bound (CRLB) for estimation of ρ_1 in a static AWGN channel is derived as $\frac{\rho_1^2}{I} + \frac{1}{2NE_{sd}^{(k)}}$. Therefore, the sampled signal ML estimator for ρ_1 in a static AWGN channel is asymptotically optimal when I is large and no decision errors occur. An unbiased estimator for ρ_1 can be derived by scaling $\hat{\rho}_1$ in (4.3) with $\frac{2I-3}{2I}$. This scaling will decrease the variance by a factor of $\frac{(2I-3)^2}{4I^2}$. This does not represent a contradiction since the ML estimator is not necessarily the optimal [14]. Moreover, when I is large, there is little difference between the performances of the original estimator and the modified estimator.

Also, from (4.2), one has

$$\hat{\rho}_2 = \hat{\rho}'_2 d \tag{4.9}$$

where $d = \frac{1}{2I-1}$ and $\hat{\rho}'_2 = \frac{R}{Q}(2I-1)$ with $R = \frac{2\hat{A}^2 E_{sd}^{(k)}}{N}$ and $Q = \frac{2I\hat{N}}{N}$. From Chapter 2, one sees that the random variable R has a normalized noncentral chi-square distribution with one degree of freedom and noncentrality parameter $\frac{2A^2 E_{sd}^{(k)}}{N}$, and the random variable Q has a normalized central chi-square distribution with $2I-1$ degrees of freedom. A singly noncentral F random variable is defined as [49, eq. (30.2)]

$$F = \frac{\chi'^2_{\nu_1}(\lambda_1)}{\chi^2_{\nu_2}} \frac{\nu_2}{\nu_1} \tag{4.10}$$

where $\chi'^2_{\nu_1}(\lambda_1)$ is a normalized noncentral chi-square variable with ν_1 degrees of freedom and noncentrality parameter λ_1, and $\chi^2_{\nu_2}$ is a normalized central chi-square variable with ν_2 degrees of freedom. Inspection of $\hat{\rho}'_2$ shows that it is a singly noncentral F random variable with parameters $\nu_1 = 1$, $\nu_2 = 2I-1$, and $\lambda_1 = \frac{2A^2 E_{sd}^{(k)}}{N}$. Then, the PDF of the sampled signal ML estimate of ρ_2 is derived as

$$p_{\hat{\rho}_2}(x) = \frac{\nu_1^{\frac{\nu_1}{2}} (d\nu_2)^{\frac{\nu_2}{2}} x^{\frac{\nu_1}{2}-1} e^{-\frac{\lambda_1}{2}}}{(d\nu_2 + \nu_1 x)^{\frac{\nu_1+\nu_2}{2}} B(\frac{\nu_1}{2}, \frac{\nu_2}{2})} \cdot {}_1F_1(\frac{\nu_1+\nu_2}{2}, \frac{\nu_1}{2}; \frac{\frac{1}{2}\lambda_1 \nu_1 x}{d\nu_2 + \nu_1 x}), \quad x \geq 0 \tag{4.11}$$

where $B(\cdot,\cdot)$ is the Beta function [2, p. 258]. The mean and the variance of $\hat{\rho}_2$ are derived from (4.11) as $E\{\hat{\rho}_2\} = \frac{2I}{2I-3} \frac{E_{sd}^{(k)}}{I} \rho_2 + \frac{1}{2I-3}$ and $Var\{\hat{\rho}_2\} = \frac{8[E_{sd}^{(k)}]^2 \rho_2^2 + 16 E_{sd}^{(k)} (I-1) \rho_2 + 4(I-1)}{(2I-3)^2 (2I-5)}$. The CRLB for estimation of ρ_2 in a static AWGN channel can also be derived as $\frac{2E_{sd}^{(k)} \rho_2}{I^2} + \frac{[E_{sd}^{(k)}]^2 \rho_2^2}{I^3}$, which includes the result in [22, eq. (64)] as a special case. Therefore, the sampled signal ML estimator for ρ_2 in a static AWGN channel is asymptotically optimal when I is large and no decision errors occur. An unbiased estimator is derived by scaling $\hat{\rho}_2$ in (4.2) with $\frac{2I-3}{2E_{sd}^{(k)}}$ and then shifting the result with $\frac{1}{2E_{sd}^{(k)}}$. The variance will vary by a factor of $\frac{(2I-3)^2}{4[E_{sd}^{(k)}]^2}$.

Results for a continuous time system can also be derived. These results will provide insights into continuous time systems using SNR estimates. Denote $\hat{\rho}_{1A} = \frac{\hat{A}}{\hat{N}_A}$ as an approximation to $\hat{\rho}_1$.

Similarly, by using (2.29) and (2.34), one has

$$p_{\hat{\rho}_{1A}}(x) = \frac{4}{\sqrt{2\pi\tilde{\sigma}^2}\Gamma(BT+\frac{1}{2})}\left(\frac{BT}{N}\right)^{BT+\frac{1}{2}} \cdot \int_0^{\infty} t^{BT+\frac{1}{2}} e^{-\frac{(x-A)^2}{2\tilde{\sigma}^2}-\frac{BT}{N}t} {}_1F_1(2, BT+\frac{1}{2}; -\frac{BT}{N}t)dt \quad (4.12)$$

where $\tilde{\sigma}^2 = \frac{N_0}{2E_{sc}^{(k)}}$ and B is the signal bandwidth defined in Section 2.1. The mean and variance of $\hat{\rho}_{1A}$ are $E\{\hat{\rho}_{1A}\} = \frac{4BTF(BT-\frac{3}{2}, BT-\frac{1}{2}; BT+\frac{1}{2}; \frac{1}{2})}{(BT-\frac{1}{2})2^{BT-\frac{1}{2}}}\rho_1$ and $Var\{\hat{\rho}_{1A}\} = E\{\hat{\rho}_{1A}^2\} - [E\{\hat{\rho}_{1A}\}]^2$, where $E\{\hat{\rho}_{1A}^2\} = \frac{4(BT)^2F(BT-\frac{3}{2}, BT-\frac{3}{2}; BT+\frac{1}{2}; \frac{1}{2})}{(BT-\frac{1}{2})(BT-\frac{3}{2})2^{BT-\frac{3}{2}}}(\frac{\tilde{\sigma}^2}{N^2} + \rho_1^2)$ and $BT > \frac{3}{2}$. They are calculated by integration from (4.12). An approximately unbiased estimator for ρ_1 is obtained by scaling $\hat{\rho}_{1A}$ with $\frac{(BT-\frac{1}{2})2^{BT-\frac{1}{2}}}{4BTF(BT-\frac{3}{2}, BT-\frac{1}{2}; BT+\frac{1}{2}; \frac{1}{2})}$. One can show that the scaling decreases the variance. Again, this is possible because the ML estimator may not be optimal, and furthermore, an approximation has been used for the continuous time PDF of $\hat{N}$.

Similarly, denote $\hat{\rho}_{2A} = \frac{E_{sc}^{(k)}}{T}\frac{\hat{A}^2}{\hat{N}_A}$ as an approximation to $\hat{\rho}_2$. One has

$$\begin{aligned} p_{\hat{\rho}_{2A}}(x) &= \frac{4\sqrt{b}e^{-\frac{A^2}{2\tilde{\sigma}^2}}}{\sqrt{2\pi\tilde{\sigma}^2 x}\Gamma(BT+\frac{1}{2})}\left(\frac{BT}{N}\right)^{BT+\frac{1}{2}} \\ &\cdot \int_0^{\infty} t^{BT} e^{-\frac{bxt}{2\tilde{\sigma}^2}-\frac{BT}{N}t} \cosh\frac{\sqrt{bxt}A}{\tilde{\sigma}^2} {}_1F_1(2, BT+\frac{1}{2}; -\frac{BT}{N}t)dt, \quad x \geq 0 \end{aligned} \quad (4.13)$$

where $b = \frac{T}{E_{sc}^{(k)}}$. The mean and variance of $\hat{\rho}_{2A}$ are $E\{\hat{\rho}_{2A}\} = \frac{4BTF(BT-\frac{3}{2}, BT-\frac{1}{2}; BT+\frac{1}{2}; \frac{1}{2})}{b(BT-\frac{1}{2})2^{BT-\frac{1}{2}}}(\rho_2 + \frac{\tilde{\sigma}^2}{N})$ and $Var\{\hat{\rho}_{2A}\} = E\{\hat{\rho}_{2A}^2\} - [E\{\hat{\rho}_{2A}\}]^2$, where $E\{\hat{\rho}_{2A}^2\} = \frac{4(BT)^2F(BT-\frac{3}{2}, BT-\frac{3}{2}; BT+\frac{1}{2}; \frac{1}{2})}{b^2(BT-\frac{1}{2})(BT-\frac{3}{2})2^{BT-\frac{3}{2}}}(\frac{3\tilde{\sigma}^4}{N^2} + \frac{6\tilde{\sigma}^2A^2}{N^2} + \rho_2^2)$ and $BT > \frac{3}{2}$. They are, again, calculated by integration from (4.13). An approximately unbiased estimator can also be obtained by scaling $\hat{\rho}_{2A}$ with $\frac{b(BT-\frac{1}{2})2^{BT-\frac{1}{2}}}{4BTF(BT-\frac{3}{2}, BT-\frac{1}{2}; BT+\frac{1}{2}; \frac{1}{2})}$ and shifting the result with $\frac{\tilde{\sigma}^2}{N}$. The variance changes accordingly.

4.1.2 Slowly Fading Channel

The ML estimators for SNRA and SNRB in a slowly fading channel can be derived in a similar manner to before. According to the invariance principle, one has

$$\hat{\rho}_1 = \hat{\rho}_1^R + j\hat{\rho}_1^I \quad (4.14)$$

and

$$\hat{\rho}_2 = c\frac{\hat{A}_R^2}{\hat{N}} + c\frac{\hat{A}_I^2}{\hat{N}} \quad (4.15)$$

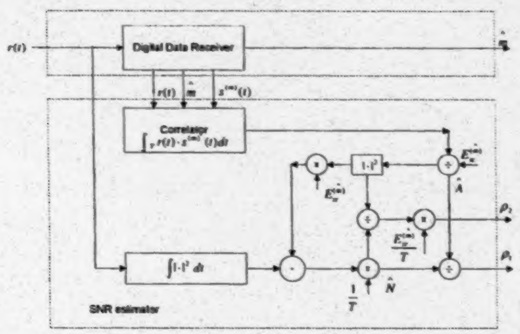
$r(t)$
Digital Data Receiver
$\hat{m}$
$r(t)$
$\hat{m}$
$s^{(\hat{m})}(t)$
Correlator
$\int_T r(t)\cdot s^{(\hat{m})*}(t)dt$
$\hat{E}_{sc}^{(\hat{m})}$
$|\cdot|^2$
$\hat{A}$
$E_{sc}^{(\hat{m})}$
ρ_2
$\frac{E_{sc}^{(\hat{m})}}{T}$
ρ_1
$\int|\cdot|^2\,dt$
$\hat{N}$
$\frac{1}{T}$
SNR estimator

Figure 4.2. Structures of the ML SNR estimators for ρ_1 and ρ_2 in a slowly fading channel.

where $\hat{\rho}_1^R = \frac{\hat{A}_R}{\hat{N}}$, $\hat{\rho}_1^I = \frac{\hat{A}_I}{\hat{N}}$, and $\hat{A}_R$ and $\hat{A}_I$ are the ML estimates of A_R and A_I in $A = A_R + jA_I$, respectively. The sampled signal ML estimators for SNRA and SNRB in slowly fading channels, $\hat{\rho}_1$ and $\hat{\rho}_2$, are derived by using equations (2.12) and (2.13) with the invariance principle as

$$\hat{\rho}_1 = \frac{I\sum_{i=1}^{I} y_i s_i^{(\hat{k})*}}{E_{sd}^{(\hat{k})}\left(\sum_{i=1}^{I}|y_i|^2\right) - \left|\sum_{i=1}^{I} y_i s_i^{(\hat{k})*}\right|^2} \tag{4.16}$$

and

$$\hat{\rho}_2 = \frac{\left|\sum_{i=1}^{I} y_i s_i^{(\hat{k})*}\right|^2}{E_{sd}^{(\hat{k})}\left(\sum_{i=1}^{I}|y_i|^2\right) - \left|\sum_{i=1}^{I} y_i s_i^{(\hat{k})*}\right|^2}. \tag{4.17}$$

The ML estimators for SNRA and SNRB in a continuous time system are obtained from (2.14) and (2.15) as

$$\hat{\rho}_1 = \frac{T\int_T y(t)s^{(\hat{k})*}(t)dt}{E_{sc}^{(\hat{k})}\left(\int_T |y(t)|^2 dt\right) - \left|\int_T y(t)s^{(\hat{k})*}(t)dt\right|^2} \tag{4.18}$$

and

$$\hat{\rho}_2 = \frac{\left|\int_T y(t)s^{(\hat{k})*}(t)dt\right|^2}{E_{sc}^{(\hat{k})}\left(\int_T |y(t)|^2\right) - \left|\int_T y(t)s^{(\hat{k})*}(t)dt\right|^2}. \tag{4.19}$$

Fig. 4.2 shows a block diagram of the ML SNRA and SNRB estimators given in (4.18) and (4.19), respectively. Similar observations can be made from Fig. 4.2. Again, this is a decision-based structure where the ML SNR estimate is based on the decision output of the digital data receiver. The difference is that, here, the correlator in Fig. 4.2 produces a complex output whereas the correlator in Fig. 4.1 produces an output that is real.

The performances of $\hat{\rho}_1$ and $\hat{\rho}_2$ in a slowly fading channel can be analyzed as previously. From (4.14), the ML estimate of ρ_1, $\hat{\rho}_1$, is completely determined by the ML estimates of ρ_1^R and ρ_1^I, $\hat{\rho}_1^R$ and $\hat{\rho}_1^I$. One sees that each of $\hat{\rho}_1^R$ and $\hat{\rho}_1^I$ has a similar form to that given in (4.1). Therefore, their PDFs are derived by solving the integral in (4.7), giving

$$p_{\hat{\rho}_1^R}(x) = \frac{I-1}{\sqrt{2\pi}x}\left(\frac{I\sigma}{Nx}\right)^{I-1} e^{\frac{1}{4}\left(\frac{I\sigma}{Nx}-\frac{A_R}{\sigma}\right)^2-\frac{A_R^2}{2\sigma^2}} D_{-I}\left(\frac{I\sigma}{Nx}-\frac{A_R}{\sigma}\right) \tag{4.20}$$

and

$$p_{\hat{\rho}_1^I}(x) = \frac{I-1}{\sqrt{2\pi}x}\left(\frac{I\sigma}{Nx}\right)^{I-1} e^{\frac{1}{4}\left(\frac{I\sigma}{Nx}-\frac{A_I}{\sigma}\right)^2-\frac{A_I^2}{2\sigma^2}} D_{-I}\left(\frac{I\sigma}{Nx}-\frac{A_I}{\sigma}\right). \tag{4.21}$$

From (4.20), one also has $E\{\hat{\rho}_1^R\} = \frac{I}{I-2}\rho_1^R$ and $Var\{\hat{\rho}_1^R\} = \frac{I^2}{(I-2)^2(I-3)}(\rho_1^R)^2 + \frac{I^2}{(I-2)(I-3)}\frac{1}{2NE_{sd}^{(k)}}$. The CRLB for estimation of ρ_1^R is $\frac{(\rho_1^R)^2}{I} + \frac{1}{2NE_{sd}^{(k)}}$. Therefore, the sampled signal ML estimator for ρ_1^R is asymptotically optimal when I is large in the absence of decision errors. An unbiased estimator can also be obtained by scaling $\hat{\rho}_1^R$ with $\frac{I-2}{I}$. The variance is reduced, as before. By inspection, similar conclusions can be made regarding $\hat{\rho}_1^I$.

Also, from (4.15), one has

$$\hat{\rho}_2 = \hat{\rho}'_2 d' \tag{4.22}$$

where $d' = \frac{1}{I-1}$ and $\hat{\rho}'_2 = \frac{R'}{Q'}(I-1)$ with $R' = \frac{2(\hat{A}_R^2+\hat{A}_I^2)E_{sd}^{(k)}}{N}$ and $Q' = \frac{2I\hat{N}}{N}$. Then, the random variable R' has a normalized noncentral chi-square distribution with two degrees of freedom and noncentrality parameter $\frac{2(A_R^2+A_I^2)E_{sd}^{(k)}}{N}$. The random variable Q' has a central chi-square distribution with $2I-2$ degrees of freedom. Therefore, $\hat{\rho}'_2$ is also a singly noncentral F random variable with parameters $v'_1 = 2$, $v'_2 = 2I-2$ and $\lambda'_1 = \frac{2(A_R^2+A_I^2)E_{sd}^{(k)}}{N}$. The PDF of the sampled signal ML estimate for ρ_2, $\hat{\rho}_2$, has a similar form to that in (4.11) except that the values of d, v_1, v_2 and λ_1 are replaced by d', v'_1, v'_2 and λ'_1, respectively. The mean and the variance of $\hat{\rho}_2$ are $E\{\hat{\rho}_2\} = \frac{I}{I-2}\frac{E_{sd}^{(k)}}{I}\rho_2 + \frac{1}{I-2}$ and $Var\{\hat{\rho}_2\} = \frac{[E_{sd}^{(k)}]^2\rho_2^2+2E_{sd}^{(k)}(I-1)\rho_2+I-1}{(I-2)^2(I-3)}$. The CRLB for estimation of ρ_2 in a slowly fading channel is $\frac{2E_{sd}^{(k)}\rho_2}{I^2} + \frac{[E_{sd}^{(k)}]^2\rho_2^2}{I^3}$. Therefore, the sampled signal ML estimator for ρ_2 in a slowly fading channel is also asymptotically optimal when I is large, in the absence of decision errors. An unbiased estimator is derived by scaling $\hat{\rho}_2$ by $\frac{I-2}{E_{sd}^{(k)}}$ and then shifting by $\frac{1}{E_{sd}^{(k)}}$. The variance changes by a factor of $\frac{(I-2)^2}{[E_{sd}^{(k)}]^2}$, as previously.

The PDFs of the approximations to the continuous time ML estimates $\hat{\rho}_1^R$ and $\hat{\rho}_1^I$, $\hat{\rho}_{1A}^R = \frac{\hat{A}_R}{\hat{N}_A}$ and $\hat{\rho}_{1A}^I = \frac{\hat{A}_I}{\hat{N}_A}$, are

$$p_{\hat{\rho}_{1A}^R}(x) = \frac{4}{\sqrt{2\pi\tilde{\sigma}^2}\Gamma(BT)} \left(\frac{BT}{N}\right)^{BT} \cdot \int_0^\infty t^{BT} e^{-\frac{(xt-A_R)^2}{2\tilde{\sigma}^2} - \frac{BT}{N}t} {}_1F_1(2, BT; -\frac{BT}{N}t)dt \qquad (4.23)$$

and

$$p_{\hat{\rho}_{1A}^I}(x) = \frac{4}{\sqrt{2\pi\tilde{\sigma}^2}\Gamma(BT)} \left(\frac{BT}{N}\right)^{BT} \cdot \int_0^\infty t^{BT} e^{-\frac{(xt-A_I)^2}{2\tilde{\sigma}^2} - \frac{BT}{N}t} {}_1F_1(2, BT; -\frac{BT}{N}t)dt. \qquad (4.24)$$

The mean and the variance of $\hat{\rho}_{1A}^R$ are $E\{\hat{\rho}_{1A}^R\} = \frac{4BTF(BT-2,BT-1;BT;\frac{1}{2})}{(BT-1)2^{BT-1}}\rho_1^R$ and $Var\{\hat{\rho}_{1A}^R\} = E\{[\hat{\rho}_{1A}^R]^2\} - \left[E\{\hat{\rho}_{1A}^R\}\right]^2$, where $E\{[\hat{\rho}_{1A}^R]^2\} = \frac{4(BT)^2F(BT-2,BT-2;BT;\frac{1}{2})}{(BT-1)(BT-2)2^{BT-2}}[\frac{\tilde{\sigma}^2}{N^2} + (\rho_1^R)^2]$ and $BT > 2$. They are calculated from (4.23) by integration. Scaling $\hat{\rho}_{1A}^R$ with $\frac{(BT-1)2^{BT-1}}{4BTF(BT-2,BT-1;BT;\frac{1}{2})}$ gives an approximately unbiased estimator. This scaling decreases the variance, as before. The performance of $\hat{\rho}_1^I$ is the same as that of $\hat{\rho}_1^R$.

The PDF of the approximation to the continuous time ML estimate $\hat{\rho}_2$, $\hat{\rho}_{2A} = \frac{E_{sc}^{(k)}}{T}\frac{\hat{A}_R^2+\hat{A}_I^2}{\hat{N}_A}$, is

$$p_{\hat{\rho}_{2A}}(x) = \frac{4e^{-\frac{|A|^2}{2\tilde{\sigma}^2}}}{2\tilde{\sigma}^2\Gamma(BT)} \left(\frac{BT}{N}\right)^{BT} \cdot \int_0^\infty t^{BT} e^{-\frac{bxt}{2\tilde{\sigma}^2} - \frac{BT}{N}t} I_0(\frac{\sqrt{bxt}|A|}{\tilde{\sigma}^2}) {}_1F_1(2, BT; -\frac{BT}{N}t)dt, \quad x \geq 0. \qquad (4.25)$$

The mean and variance of $\hat{\rho}_{2A}$ are $E\{\hat{\rho}_{2A}\} = \frac{4BTF(BT-2,BT-1;BT;\frac{1}{2})}{m(BT-1)2^{BT-1}}(\rho_2 + \frac{2\tilde{\sigma}^2}{N})$ and $Var\{\hat{\rho}_{2A}\} = E\{\hat{\rho}_{2A}^2\} - [E\{\hat{\rho}_{2A}\}]^2$, where $E\{\hat{\rho}_{2A}^2\} = \frac{4(BT)^2F(BT-2,BT-2;BT;\frac{1}{2})}{b^2(BT-1)(BT-2)2^{BT-2}}(\rho_2^2 + \frac{8\tilde{\sigma}^4}{N^2} + \frac{8\tilde{\sigma}^2|A|^2}{N^2})$ and $BT > 2$. They are also calculated by integration from (4.25). An approximately unbiased estimator is obtained by scaling $\hat{\rho}_{2A}$ with $\frac{b(BT-1)2^{BT-1}}{4BTF(BT-2,BT-1;BT;\frac{1}{2})}$ and shifting the result with $\frac{2\tilde{\sigma}^2}{N}$. The variance changes as well.

4.1.3 Numerical Examples

Since the SNRA estimator in a slowly fading channel is completely determined by $\hat{\rho}_1^R$ and $\hat{\rho}_1^I$ while $\hat{\rho}_1^R$ and $\hat{\rho}_1^I$ have the same performance, only $\hat{\rho}_1^R$ is examined. Figs. 4.3 and 4.4 show the biases and the root mean squared errors (RMSEs) of the sampled signal SNRA estimators, where $\frac{1}{2NE_{sd}^{(k)}}$ is set equal to 1. As can be seen from Fig. 4.3, at $I = 8$, $\hat{\rho}_1$ in an AWGN channel has a bias of about 1.80 dB and $\hat{\rho}_1^R$ in a slowly fading channel has a bias of about 2.50 dB, while at $I = 20$, $\hat{\rho}_1$ has a bias of about 0.70 dB and $\hat{\rho}_1^R$ has a bias of about 0.90 dB. Thus, a larger value of I corresponds to a smaller bias of the estimator. One can also see from Fig. 4.4 that the RMSEs of the estimators decrease as I increases. Therefore, the performance of the estimator in both channels is improved by increasing I, as expected. Figs. 4.5 and 4.6 show the performances of the continuous time signal SNRA estimators, where $\frac{\tilde{\sigma}^2}{N^2}$ is set equal to 1. Similar observations to those made from Figs. 4.3 and 4.4 can be made, with BT in Figs. 4.5 and 4.6 playing a similar role to the role that I plays in Figs. 4.3 and 4.4. Comparing the continuous time signal estimators with the sampled signal estimators, one sees that the continuous time signal estimators perform slightly better than the corresponding sampled signal estimators. As an example, at $BT = 8$ and $\rho_1^R = 60\ dB$, the continuous time signal estimator for ρ_1^R in a slowly fading channel has a bias of about 2.4 dB and a RMSE of about 630 while the sampled signal estimator for ρ_1^R in a slowly fading channel has a bias of about 2.5 dB and a RMSE of about 690.

Figs. 4.7 and 4.8 show the performances of the sampled signal SNRB estimators, where the sampled signal power $\frac{E_{sd}^{(k)}}{I}$ is set equal to 1. Again, the biases and RMSEs of the estimators decrease as I increases. Note that $\hat{\rho}_2$ has a fairly large bias at small values of ρ_2, as $E\{\hat{\rho}_2\}$ has a shifting factor, which is dominant when ρ_2 is small. Figs. 4.9 and 4.10 show the performances of the continuous time signal ML estimators for SNRB, where the continuous time signal power $\frac{E_{sc}^{(k)}}{T}$ is set equal to 1. Again, the continuous time signal estimators perform slightly better than the corresponding sampled signal estimators. One concludes from Figs. 4.3 to 4.10 that the SNR estimator in an AWGN channel outperforms that in a slowly fading channel.

Figs. 4.11 and 4.12 compare the PDFs of the sampled signal ML estimates of ρ_1 and ρ_2 in an AWGN channel with the limiting Gaussian PDF's. The limiting Gaussian PDF has a variance equal to the CRLB and a mean equal to the true value of ρ_1 or ρ_2, which has been fixed to 10. One sees that the difference between the estimate PDF and the Gaussian PDF decreases as I increases, as

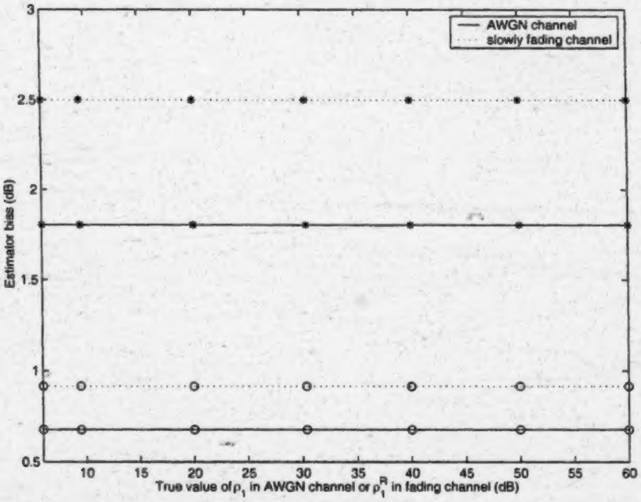

Figure 4.3. The biases of the sampled signal ML estimator for ρ_1 in an AWGN channel and ρ_1^R in a slowly fading channel at $I = 8$ (*) and $I = 20$ (o).

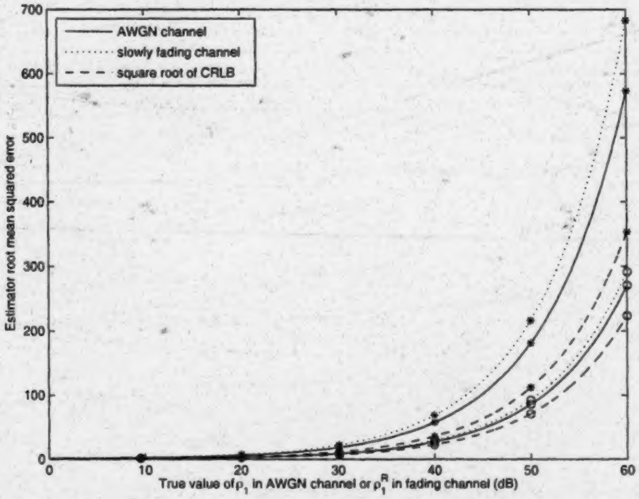

Figure 4.4. The RMSEs of the sampled signal ML estimator for ρ_1 in an AWGN channel and ρ_1^R in a slowly fading channel at $I = 8$ (*) and $I = 20$ (o).

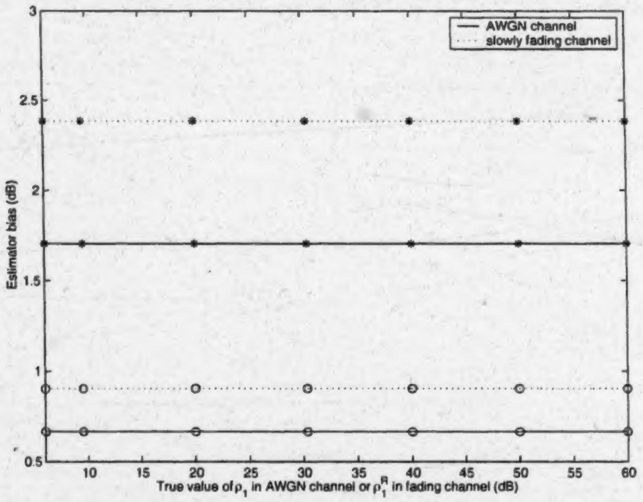

Figure 4.5. The biases of the continuous time signal ML estimator for ρ_1 in an AWGN channel and ρ_1^R in a slowly fading channel at $BT = 8$ (*) and $BT = 20$ (o).

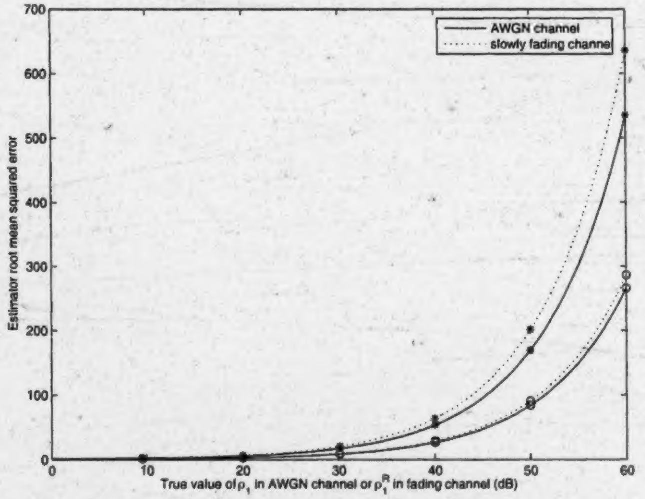

Figure 4.6. The RMSEs of the continuous time signal ML estimator for ρ_1 in an AWGN channel and ρ_1^R in a slowly fading channel at $BT = 8$ (*) and $BT = 20$ (o).

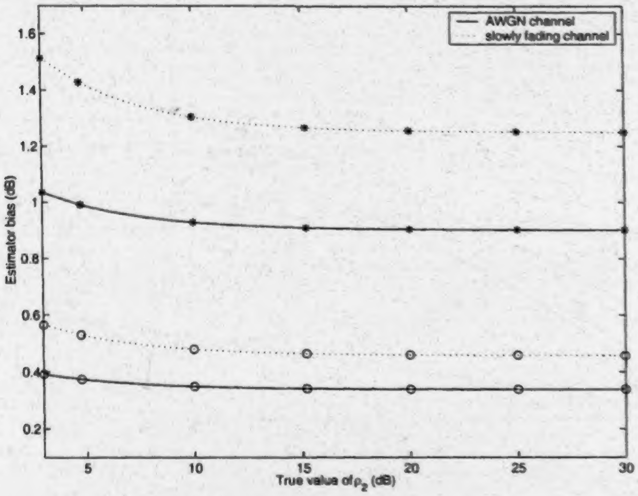

Figure 4.7. The biases of the sampled signal ML estimator for ρ_2 in an AWGN channel and a slowly fading channel at $I = 8$ (*) and $I = 20$ (o).

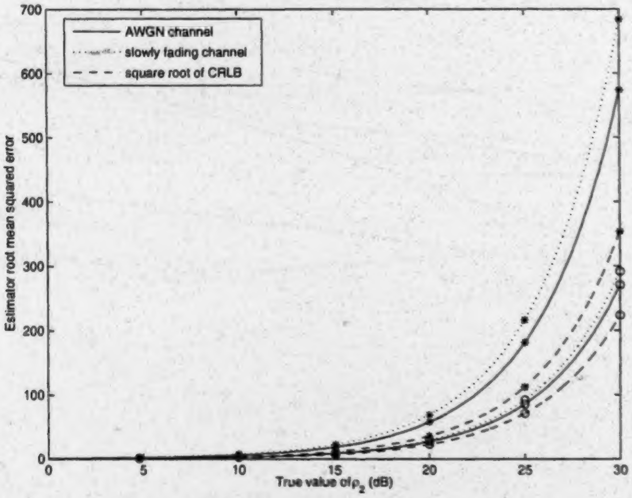

Figure 4.8. The RMSEs of the sampled signal ML estimator for ρ_2 in an AWGN channel and a slowly fading channel at $I = 8$ (*) and $I = 20$ (o).

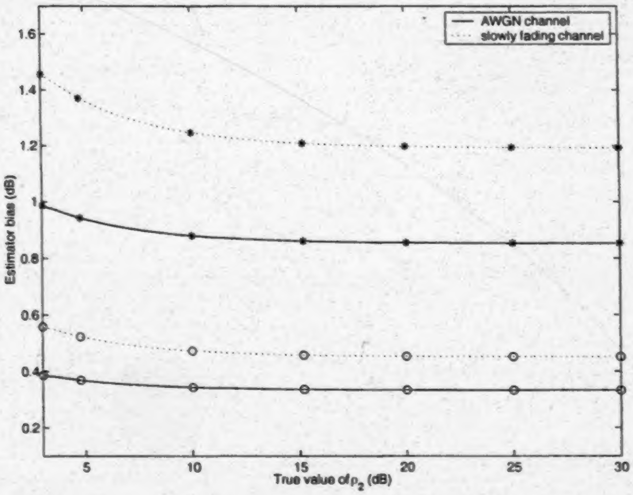

Figure 4.9. The biases of the continuous time signal ML estimator for ρ_2 in an AWGN channel and a slowly fading channel at $BT = 8$ (*) and $BT = 20$ (o).

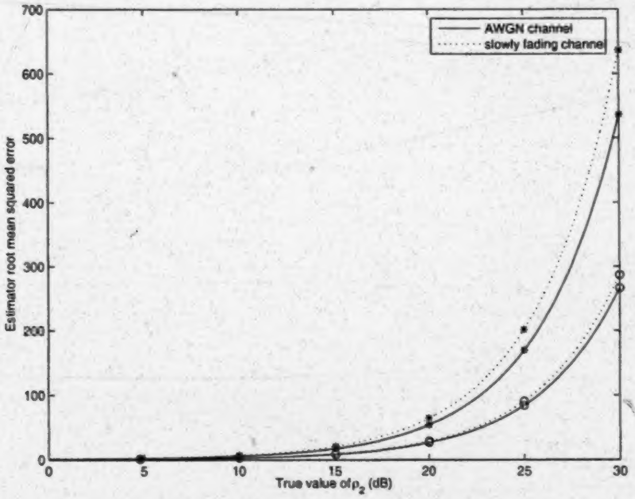

Figure 4.10. The RMSEs of the continuous time signal ML estimator for ρ_2 in an AWGN channel and a slowly fading channel at $BT = 8$ (*) and $BT = 20$ (o).

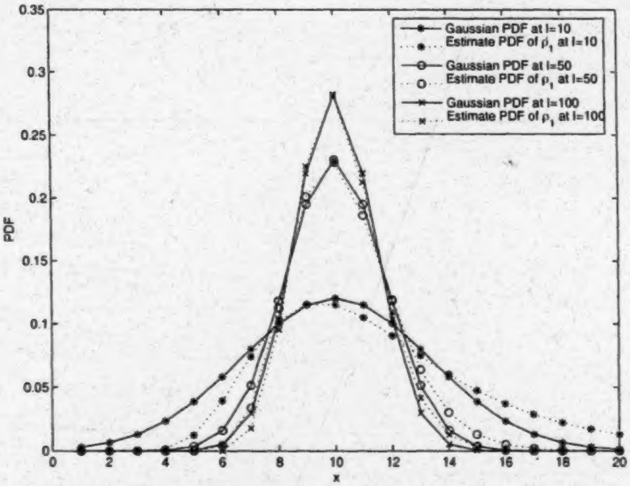

Figure 4.11. Comparison of the limiting Gaussian PDF and the PDF of the sampled signal ML estimator $\hat{\rho}_1$ in an AWGN channel for different values of sample sizes.

expected [15]. At $I = 100$, the estimate PDF can be approximated by the Gaussian PDF graphically. The approximation errors around $x = 7$ and $x = 13$ in Figs. 4.11 and 4.12 won't diminish until I goes to infinity and both the estimate PDF and the Gaussian PDF become pulse functions. The PDFs of the sampled signal ML estimates in a slowly fading channel and the continuous time signal ML estimates can be examined in a similar way.

The above results are based on the assumption of no decision errors. They are valid for data-aided (DA) estimation. In the case when decision-based estimation is performed, they are also valid for moderate to large values of SNR. In applications such as rate adaptation and power control, the SNR estimate is used to control the transmission rate and power of subsequent symbols. If the value of SNR in the decision-based estimation is small, by a proper receiver design, the error rate can still be kept low. Moreover, as can be seen from the numerical results, good estimator performances are achieved with tens of samples. Even an error rate as large as 10^{-2} still implies that the decision-based estimation is probably accomplished in the absence of decision errors. In this case, the results

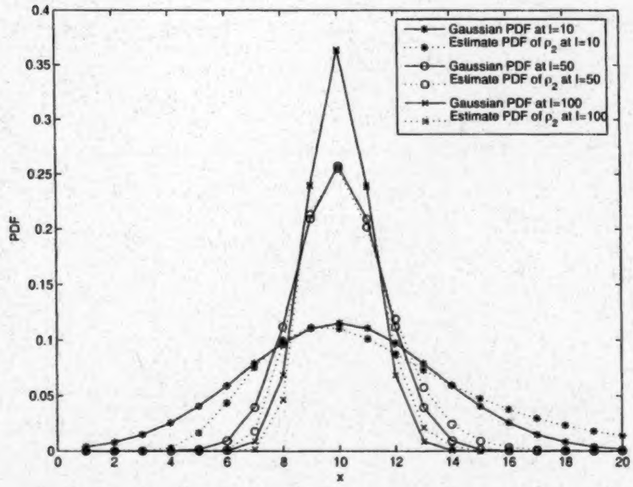

Figure 4.12. Comparison of the limiting Gaussian PDF and the PDF of the sampled signal ML estimator $\hat{\rho}_2$ in an AWGN channel for different values of sample sizes.

obtained under an error-free operation assumption are useful [20], [22].

4.2 ML Estimation of SNR Using Both Pilot and Data Symbols

The previous SNR estimators use either known or unknown symbols. In a practical communication system, a frame usually consists of both known pilot symbols and unknown data symbols. It is advantageous to use all available symbols in the frame to estimate the SNR as accurately as possible. In this work, we derive a novel ML-based SNR estimator for binary phase shift keying (BPSK) signals using both pilot and data symbols simultaneously. The obtained estimator is simple. It can be applied to systems where signals are transmitted in frames with both known and unknown fields.

4.2.1 Estimator Structure

Using the same system model as those in [19]- [27], one has the received signals after matched filtering as

$$y_l = As_l + n_l \tag{4.26}$$

where $l = 1,2,\cdots,L+Q$ index different symbols in the frame, A is the channel gain assumed constant over the frame, $s_l \in \{-1,1\}$ is the transmitted BPSK signal of the l-th symbol, and n_l is the noise term of the l-th symbol assumed to be a Gaussian random variable with mean 0 and variance σ^2. Without loss of generality, assume that $l = 1,2,\cdots,L$ index the pilot symbols in the frame and that $l = L+1, L+2,\cdots,L+Q$ index the data symbols in the frame. Also, assume $s_l = 1$ ($l = 1,2,\cdots,L$) as the pilot symbols. The SNR to be estimated is defined as $\rho = \frac{A^2}{\sigma^2}$. The probability density function (PDF) of the received signals of the pilot symbols can be derived from (4.26) as

$$f(y_l;A,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_l-A)^2}{2\sigma^2}} \tag{4.27}$$

where $l = 1,2,\cdots,L$. Assuming that the BPSK signals have equal *a priori* probabilities, the PDF of the received signals of the data symbols is derived from (4.26) as

$$f(y_l;A,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y_l^2+A^2}{2\sigma^2}} \cosh\left(\frac{y_l A}{\sigma^2}\right) \tag{4.28}$$

where $l = L+1, L+2,\cdots,L+Q$ and $\cosh(\cdot)$ is the hyperbolic cosine function. Finally, assuming n_l ($l = 1,2,\cdots,L+Q$) are independent, one has from (4.27) and (4.28) the log-likelihood function

$$\ln f(\mathbf{y};A,\rho) = \frac{L+Q}{2}\ln\left(\frac{\rho}{2\pi A^2}\right) - \frac{\rho\sum_{l=1}^{L+Q} y_l^2}{2A^2} - \frac{(L+Q)\rho}{2} + \frac{\rho\sum_{l=1}^{L} y_l}{A} + \sum_{l=L+1}^{L+Q} \ln\cosh\left(\frac{y_l\rho}{A}\right) \tag{4.29}$$

where $\mathbf{y} = [y_1, \quad y_2, \quad \cdots, \quad y_{L+Q}]$ and $\rho = \frac{A^2}{\sigma^2}$ has been used. By differentiating (4.29) with respect to A and ρ and setting the derivatives equal to zero, one derives the two equations

$$\rho\sum_{l=1}^{L+Q} y_l^2 - (L+Q)A^2 - A\rho\sum_{l=1}^{L} y_l - A\rho\sum_{l=L+1}^{L+Q} \tanh\left(\frac{y_l\rho}{A}\right)y_l = 0 \tag{4.30}$$

and

$$\rho\sum_{l=1}^{L+Q} y_l^2 + (L+Q)(\rho-1)A^2 - 2A\rho\sum_{l=1}^{L} y_l - 2A\rho\sum_{l=L+1}^{L+Q} \tanh\left(\frac{y_l\rho}{A}\right)y_l = 0 \tag{4.31}$$

where $\tanh(\cdot)$ is the hyperbolic tangent function. An exact solution of (4.30) and (4.31) for ρ is difficult to obtain due to the nonlinearity of the hyperbolic tangent function. Note that $\tanh\left(\frac{y_l\rho}{A}\right) \approx +1$ when $\frac{y_l\rho}{A} > 0$ and $\tanh\left(\frac{y_l\rho}{A}\right) \approx -1$ when $\frac{y_l\rho}{A} < 0$, for sufficiently large ρ [25]. Using this approximation, one can solve (4.30) and (4.31) to give an approximate ML-based estimator for ρ as

$$\hat{\rho}' \approx \frac{\hat{A}^2}{\frac{1}{L+Q}\sum_{l=1}^{L+Q} y_l^2 - \hat{A}^2} \tag{4.32}$$

where $\hat{A} = \frac{1}{L+Q}\left(\sum_{l=1}^{L} y_l + \sum_{l=L+1}^{L+Q} |y_l|\right)$. The denominator of (4.32) is actually a biased estimate of σ^2 [93]. A reduced-bias ML-based estimator for ρ can be derived by multiplying the denominator of (4.32) with $\frac{L+Q}{L+Q-1}$, giving

$$\hat{\rho} \approx \frac{\hat{A}^2}{\frac{1}{L+Q-1}\sum_{l=1}^{L+Q} y_l^2 - \frac{L+Q}{L+Q-1}\hat{A}^2}. \tag{4.33}$$

It can be shown that the CRLB for estimation of ρ when L pilot symbols and Q data symbols are available is $\frac{2\rho(\rho+2)(L+Q)-4\rho^2 Qc}{(L+Q)^2-2Qc-2\rho Q^2c^2+4(L+Q)Qc}$, where $c = \sqrt{\frac{2}{\pi\rho}}e^{-\frac{\rho}{2}} + erf\left(\sqrt{\frac{\rho}{2}}\right) - 1$ and $erf(\cdot)$ is the error function. Note that the SNR estimator in (4.33) uses the received signals of both the pilot symbols and the data symbols in the frame. Two special cases are considered below. If one uses the received signals of the pilot symbols in the frame only, $\hat{\rho}$ becomes

$$\hat{\rho}_1 = \frac{\hat{A}_1^2}{\frac{1}{L-1}\sum_{l=1}^{L} y_l^2 - \frac{L}{L-1}\hat{A}_1^2} \tag{4.34}$$

where $\hat{A}_1 = \frac{1}{L}\sum_{l=1}^{L} y_l$. On the other hand, if one uses the received signals of the data symbols in the frame only, $\hat{\rho}$ becomes

$$\hat{\rho}_2 \approx \frac{\hat{A}_2^2}{\frac{1}{Q-1}\sum_{l=L+1}^{L+Q} y_l^2 - \frac{Q}{Q-1}\hat{A}_2^2} \tag{4.35}$$

where $\hat{A}_2 = \frac{1}{Q}\sum_{l=L+1}^{L+Q} |y_l|$ and an approximation similar to that in (4.33) is also used. Observe that $\hat{A}$ is a linear combination of $\hat{A}_1$ and $\hat{A}_2$, as $\hat{A} = \frac{L}{L+Q}\hat{A}_1 + \frac{Q}{L+Q}\hat{A}_2$, but $\hat{\rho}$ is not a linear combination of $\hat{\rho}_1$ and $\hat{\rho}_2$.

4.2.2 Performance Analysis

Figs. 4.13 and 4.14 compare the performances of $\hat{\rho}$, $\hat{\rho}_1$ and $\hat{\rho}_2$ when $L = 8$ and $Q = 28$. One observes that, when $SNR \geq 9.0\ dB$, the bias and the normalized root mean square error (NRMSE)

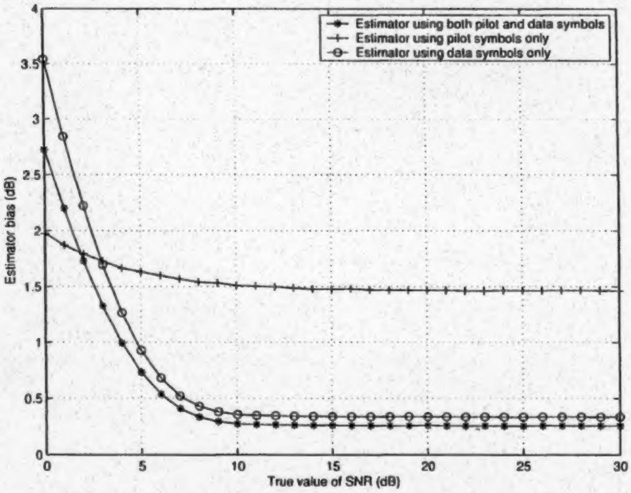

Figure 4.13. The biases of $\hat{\rho}$ (using both pilot symbols and data symbols), $\hat{\rho}_1$ (using pilot symbols only), and $\hat{\rho}_2$ (using data symbols only), when $L = 8$ and $Q = 28$.

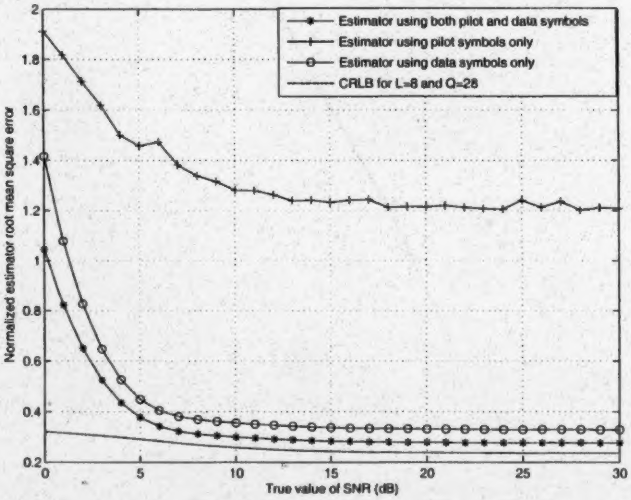

Figure 4.14. The normalized root mean squared errors of $\hat{\rho}$ (using both pilot symbols and data symbols), $\hat{\rho}_1$ (using pilot symbols only), and $\hat{\rho}_2$ (using data symbols only), when $L = 8$ and $Q = 28$, and the Cramér-Rao lower bound.

of $\hat{\rho}$ are approximately constant at 0.25 dB and 0.30, respectively, and the estimator $\hat{\rho}$ performs well. It performs poorly, however, when the SNR is less than about 2.0 dB. Since both $\hat{\rho}$ and $\hat{\rho}_2$ use the approximation of the $\tanh(\cdot)$ function in their derivations and their performances demonstrate similar behaviors, as can be seen from Figs. 4.13 and 4.14, the poor performance of $\hat{\rho}$ at small values of SNR is probably caused by the approximation errors in (4.33). (Observe further that $\hat{\rho}_1$ does not use this approximation and its performance degradation with decreasing SNR is not as severe as is the case for $\hat{\rho}$ and $\hat{\rho}_2$). Comparing $\hat{\rho}$ with $\hat{\rho}_1$ and $\hat{\rho}_2$, one observes that $\hat{\rho}$ outperforms $\hat{\rho}_1$ for most values of SNR considered, and outperforms $\hat{\rho}_2$ for all the values of SNR considered. When the SNR is less than about 1.9 dB, $\hat{\rho}$ has a larger bias than $\hat{\rho}_1$. This is explained as follows. First, from (4.33) and (4.34), the main difference between $\hat{\gamma}$ and $\hat{\rho}_1$ is the fact that $\hat{\rho}$ uses the received signals of Q unknown data symbols, in addition to the received signals of L known pilot symbols used by $\hat{\rho}_1$. This leads to two consequences. On the one hand, $\hat{\rho}$ benefits from a larger sample size by including the data symbols in the estimation. On the other hand, $\hat{\rho}$ also suffers from the approximation error introduced by using these unknown symbols. As a result, one sees in Fig. 4.13 that the performance gain of $\hat{\rho}$ over $\hat{\rho}_1$ for a fixed Q decreases as the SNR decreases from 9.0 dB to 2.0 dB, since the approximation error cannot be ignored in this case and it increases as the SNR decreases. At SNR values less than about 1.9 dB, the penalty incurred by using the data symbols is too large to be compensated by the gain, and $\hat{\rho}$ shows a larger bias than $\hat{\rho}_1$ in this SNR region. Second, from (4.33) and (4.35), the main difference between $\hat{\rho}$ and $\hat{\rho}_2$ is the fact that $\hat{\rho}$ uses the received signals of L known pilot symbols, in addition to the received signals of M unknown data symbols used by $\hat{\rho}_2$. Unlike the previous case, since the pilot symbols are known, there is no penalty incurred from using the pilot symbols. As a result, $\hat{\rho}$ outperforms $\hat{\rho}_2$ for all values of SNR considered, as seen in Figs. 4.13 and 4.14, as it always benefits from a larger sample size by including the pilot symbols in the estimation.

Figs. 4.15 and 4.16 compare the performances of $\hat{\rho}$, the M_2M_4 estimator in [22, eq. (42)] and the SNV estimator in [22, eq. (30)] when $L = 8$ and $Q = 28$. To make a fair comparison, we assume that both the pilot symbols and the data symbols in the frame are available to the M_2M_4

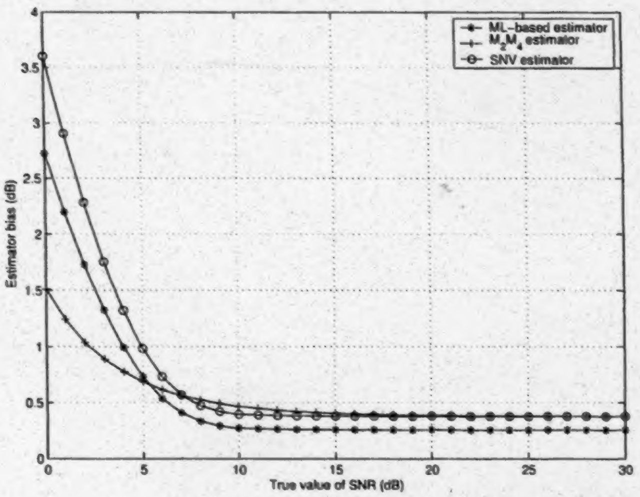

Figure 4.15. The biases of the ML-based estimator $\hat{\rho}$, the M_2M_4 estimator and the SNV estimator when $L = 8$ and $Q = 28$.

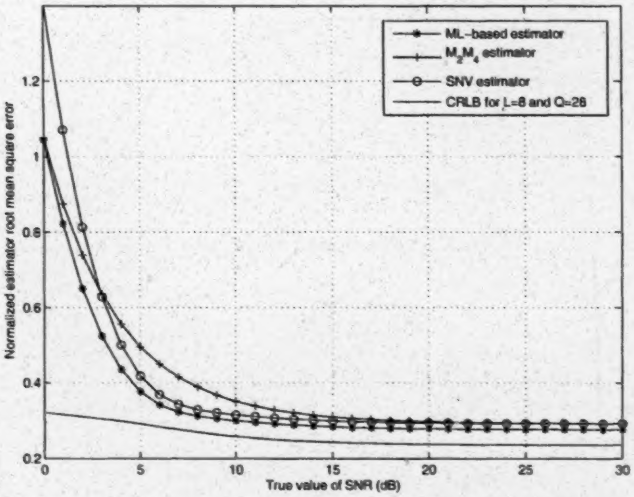

Figure 4.16. The normalized root mean squared errors of the ML-based estimator $\hat{\rho}$, the M_2M_4 estimator and the SNV estimator when $L = 8$ and $Q = 28$, and the Cramér-Rao lower bound.

estimator and the SNV estimator. One sees from Figs. 4.15 and 4.16 that $\hat{\rho}$ performs better than the SNV estimator for all values of SNR considered. As an example, $\hat{\rho}$ always has a bias gain of about 0.1 dB over the SNV estimator at the considered values of SNR. Comparing $\hat{\rho}$ with the M_2M_4 estimator, one observes that $\hat{\rho}$ has smaller bias and NRMSE than the M_2M_4 estimator, when $30.0\,dB \geq SNR \geq 5.0\,dB$. Therefore, $\hat{\rho}$ performs better than the M_2M_4 estimator when $30.0\,dB \geq SNR \geq 5.0\,dB$. When $5.0\,dB \geq SNR \geq 0.0\,dB$, $\hat{\rho}$ has a smaller NRMSE but a larger bias than the M_2M_4 estimator. The performance of the M_2M_4 estimator is mainly determined by errors in the approximation of the second- and fourth-order moments [22], while the performance of $\hat{\rho}$ is mainly determined by errors in the approximation of the $\tanh(\cdot)$ function in this case. From Figs. 4.15 and 4.16, the performance of $\hat{\rho}$ is more sensitive to the value of SNR than the performance of the M_2M_4 estimator. This suggests that the approximation error in the $\tanh(\cdot)$ function used to derive the ML-based estimator $\hat{\rho}$ is more sensitive to the value of SNR than the approximation errors in the second- and fourth-order moments used to derive the M_2M_4 estimator. As a result, increasing the SNR improves the performance of $\hat{\rho}$ greatly. It can also be shown that $\hat{\rho}$ outperforms the SNV estimator over the SNR region of interest, while $\hat{\rho}$ outperforms the M_2M_4 estimator at large values of SNR and underperforms the M_2M_4 estimator at small values of SNR, when other values of L and Q are used. The preceding estimator is obtained for BPSK signaling, which is used, for example, in a cdma2000 system [94].

4.3 ML Estimation of ASNR in a Ricean Channel

In the previous two sections, we have actually estimated the instantaneous SNR. In this work, we derive ML estimator for ASNR in a slowly Ricean fading channel.

4.3.1 Estimator Structure

Use the same system model as that in Section 3.2. The i-th component of the received signal is given by (3.18), where r_l and θ_l here are the Ricean fading envelope and the Ricean fading phase, respectively. The received signal is correlated with the known transmitted signal, giving

$$w_l = \sum_{i=1}^{I} y_{il} s_{il}^{*} = r_l e^{j\theta_l} E + n_l \tag{4.36}$$

where $E = \sum_{i=1}^{I} |s_{il}|^2$ is the signal energy assuming all signals used in the estimation have the same energy, and $n_l = \sum_{i=1}^{I} n_{il} s_{il}^{*}$ is a complex Gaussian random variable with mean 0 and variance $E\{n_l n_l^{*}\} = N_0 E$. The instantaneous SNR can be derived from (4.36) as $\frac{E}{N_0} r_l^2$. Note that it is changing from symbol to symbol since r_l is a random variable. A useful performance measure is the ASNR, defined as

$$\bar{\rho} = \frac{E}{N_0} E\{r_l^2\} = \frac{\Omega}{2\sigma^2} \tag{4.37}$$

where $\Omega = E\{r_l^2\} = P^2 + 2\alpha^2$ is the total mean power of the fading signal, P^2 is the mean power of the LOS component, $2\alpha^2$ is the mean power of scattering component, and $2\sigma^2 = \frac{N_0}{E}$.

From the invariance principle of ML estimation, one has that the ML estimator for $\bar{\rho}$, $\hat{\bar{\rho}}$, satisfies

$$\hat{\bar{\rho}} = \frac{\hat{\Omega}}{\widehat{2\sigma^2}} \tag{4.38}$$

where $\hat{\Omega}$ and $\widehat{2\sigma^2}$ are the ML estimates of Ω and $2\sigma^2$, respectively. It has been shown in Section 3.2 that the ML estimators for $2\sigma^2$ and Ω in a Ricean fading channel, $\widehat{2\sigma^2}$ and $\hat{\Omega}$, are

$$\widehat{2\sigma^2} = \frac{1}{(I-1)LE} \sum_{l=1}^{L} \sum_{i=1}^{I} |y_{il}|^2 - \frac{1}{(I-1)L} \sum_{l=1}^{L} z_l^2 \tag{4.39}$$

and

$$\hat{\Omega} = \frac{1}{L} \sum_{l=1}^{L} z_l^2 - \widehat{2\sigma^2} \tag{4.40}$$

where $z_l = \frac{1}{E}|w_l|$ is the normalized absolute value of the correlator output in (4.36) and is also the ML estimate of r_i as shown in Section 2.1. Using (4.38), (4.39) and (4.40), $\hat{\bar{\rho}}$ is then

$$\hat{\bar{\rho}} = \frac{X}{Y} - 1 \tag{4.41}$$

where $X = \frac{1}{L} \sum_{l=1}^{L} z_l^2$ and $Y = \frac{1}{(I-1)LE} \sum_{l=1}^{L} \sum_{i=1}^{I} |y_{il}|^2 - \frac{1}{(I-1)L} \sum_{l=1}^{L} z_l^2$. By taking an average of the ML estimates of the instantaneous SNR in L symbol intervals, one can also obtain a moment-based estimator for $\bar{\rho}$ as $\hat{\bar{\rho}}_M = \frac{1}{L} \sum_{l=1}^{L} \hat{\bar{\rho}}_l$, where $\hat{\bar{\rho}}_l = \frac{z_l^2}{\frac{1}{I} \sum_{i=1}^{I} |y_{il}|^2 - \frac{E}{I} z_l^2}$ is the ML estimate of the instantaneous

SNR using the l-th symbol only [22]. It can be shown that $\hat{\rho}_M$ has larger bias, variance and mean square error than $\hat{\rho}$. Therefore, the ML estimator, $\hat{\rho}$, outperforms the moment-based estimator, $\hat{\rho}_M$, as expected. Next, we examine the performance of the ML estimator, $\hat{\rho}$, analytically.

4.3.2 Analytical Performance

We derive the PDFs of X and Y first. It was derived in Section 3.1 that the PDF of z_l in a noisy Ricean fading channel is given by (3.5). Since $X = \frac{1}{L}\sum_{l=1}^{L} z_l^2$ is a sum of z_l^2, $l = 1, 2, \cdots, L$, one has the PDF of X [1, eq. (2-1-118)]

$$p_X(x) = \frac{1}{2\lambda_x}\left(\frac{x}{P^2}\right)^{\frac{L-1}{2}} e^{-\frac{P^2+x}{2\lambda_x}} I_{L-1}\left(\frac{\sqrt{xP^2}}{\lambda_x}\right), \quad x \geq 0 \tag{4.42}$$

which is a noncentral chi-square distribution with $2L$ degrees of freedom, noncentrality parameter P^2, and parameter $\lambda_x = \frac{1}{L}(\sigma^2 + \alpha^2)$. The PDF of Y can be derived from (4.39) as a central chi-square distribution with $2(I-1)L$ degrees of freedom and parameter $\lambda_y = \frac{2\sigma^2}{2(I-1)L}$, which is given by [1, eq. (2-1-110)]

$$p_Y(y) = \frac{1}{(2\lambda_y)^{(I-1)L}\Gamma((I-1)L)} y^{(I-1)L-1} e^{-\frac{y}{2\lambda_y}}, \quad y \geq 0. \tag{4.43}$$

Moreover, as discussed in Chapter 2, Y is independent of X.

Using (4.42) and (4.43), together with (4.41), the PDF of $\hat{\rho}$ can then be derived. Denote $X' = \frac{X}{\lambda_x}$ as a normalized noncentral chi-square random variable with $2L$ degrees of freedom and noncentrality parameter $\frac{LP^2}{\sigma^2+\alpha^2}$, and $Y' = \frac{Y}{\lambda_y}$ as a normalized central chi-square variable with $2(I-1)L$ degrees of freedom. Also, rewrite (4.41) as

$$\hat{\rho} = \bar{\rho}' \cdot \frac{2\alpha^2 + 2\sigma^2}{2\sigma^2} - 1 \tag{4.44a}$$

$$\bar{\rho}' = \frac{X'}{2L}\left[\frac{Y'}{2(I-1)L}\right]^{-1}. \tag{4.44b}$$

One sees that $\bar{\rho}'$ is a singly noncentral F-variable with parameters $\nu_1 = 2L$, $\nu_2 = 2(I-1)L$ and $\lambda_1 = \frac{LP^2}{\sigma^2+\alpha^2}$. From [49, eq. (30.8)], $\hat{\rho}$ has PDF

$$p(\hat{\rho}) \;=\; \frac{(2\sigma^2\nu_1)^{\frac{\nu_1}{2}}(2\sigma^2\nu_2 + 2\alpha^2\nu_2)^{\frac{\nu_2}{2}}}{B(\frac{\nu_1}{2}, \frac{\nu_2}{2})(2\sigma^2\nu_1 + 2\sigma^2\nu_2 + 2\alpha^2\nu_2)^{\frac{\nu_1+\nu_2}{2}}} \cdot \frac{(1+\hat{\rho})^{\frac{\nu_1}{2}-1} e^{-\frac{\lambda_1}{2}}}{\left(\frac{2\sigma^2\nu_1}{2\sigma^2\nu_1+2\sigma^2\nu_2+2\alpha^2\nu_2}\hat{\rho} + 1\right)^{\frac{\nu_1+\nu_2}{2}}} \tag{4.45}$$

$$\cdot {}_1F_1\left(\frac{\nu_1+\nu_2}{2}, \frac{\nu_1}{2}; \frac{\frac{1}{2}\lambda_1\nu_1(1+\hat{\bar{\rho}})\sigma^2}{\sigma^2\nu_1(1+\hat{\bar{\rho}})+\sigma^2\nu_2+\alpha^2\nu_2}\right)$$

where ${}_1F_1(\cdot,\cdot;\cdot)$ is the confluent hypergeometric function [2, p. 504]. The mean, the variance and the mean square error of $\hat{\bar{\rho}}$ are $E\{\hat{\bar{\rho}}\} = \frac{(I-1)L}{(I-1)L-1}\bar{\rho} + \frac{1}{(I-1)L-1}$, $Var\{\hat{\bar{\rho}}\} = \frac{[(I-1)L]^2(LI-1)}{[(I-1)L-1]^2[(I-1)L-2]} \cdot \frac{(\bar{\rho}+1)^2}{L} - \frac{[(I-1)L]^2}{[(I-1)L-1][(I-1)L-2]} \cdot \frac{1}{L}(\frac{K}{K+1})^2\bar{\rho}^2$ and $MSE\{\hat{\bar{\rho}}\} = Var\{\hat{\bar{\rho}}\} + (E\{\hat{\bar{\rho}}\} - \bar{\rho})^2$ as $P^2 = \frac{K}{K+1}\Omega$. The Cramér-Rao lower bound for estimation of $\bar{\rho}$ in a Ricean fading channel is calculated numerically. Figs. 4.17 to 4.19 show the bias, the normalized variance and the normalized mean square error of $\hat{\bar{\rho}}$. One sees that even when $L = 100$, the bias of $\hat{\bar{\rho}}$ is less than 0.09 dB, the normalized variance of $\hat{\bar{\rho}}$ deviates from the CRLB within a range of 3% to 4%, and the normalized mean square error of $\hat{\bar{\rho}}$ deviates from the CRLB within a range of 4% to 5%. As the sample size increases, the bias and the deviations are further reduced. An unbiased estimator for $\bar{\rho}$ can be obtained by shifting $\hat{\bar{\rho}}$ with $\frac{1}{(I-1)L-1}$ and then scaling with $\frac{(I-1)L-1}{(I-1)L}$, giving

$$\tilde{\bar{\rho}} = \frac{(I-1)L-1}{(I-1)L}\left[\hat{\bar{\rho}} - \frac{1}{(I-1)L-1}\right]. \tag{4.46}$$

One also has $Var\{\tilde{\bar{\rho}}\} = MSE\{\hat{\bar{\rho}}\} = \frac{LI-1}{(I-1)L-2}\frac{(\bar{\rho}+1)^2}{L} - \frac{(I-1)L-1}{(I-1)L-2}\frac{1}{L}(\frac{K}{K+1})^2\bar{\rho}^2$, which are smaller than the corresponding values of $\hat{\bar{\rho}}$. Therefore, the estimator $\tilde{\bar{\rho}}$ has better performance than $\hat{\bar{\rho}}$.

4.4 Joint Estimation of ASNR and K Parameter in a Ricean Channel

In the previous section, estimation of ASNR for a noisy Ricean fading channel has been studied. As well, in Chapter 3, estimators for the Ricean K parameter using noisy samples have been proposed. These estimators estimate the ASNR and the Ricean K parameter separately by using independent samples. In some cases, it is desirable to perform a joint estimation of the ASNR and the Ricean K parameter. Also, in a practical wireless communication system, the channel samples are usually correlated. As an example, in a communication system with a carrier frequency of 900 MHz, if the mobile moves at a speed of 100 Km/h and the received signal is sampled at a rate of 2 KHz, the covariance between the neighboring two channel samples is about 0.9829, according to Jakes' model [54]. Therefore, it is of great practical interest to design estimators for the Ricean K parameter

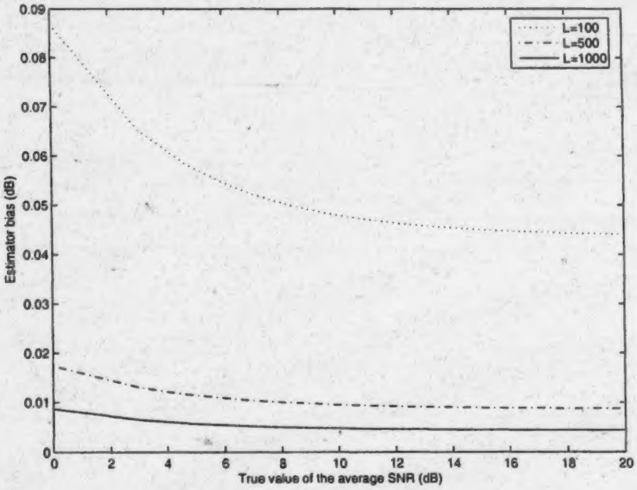

Figure 4.17. The biases (in dB) of $\hat{\bar{\rho}}$ for $I = 2$ and $K = 5$ for different sample sizes.

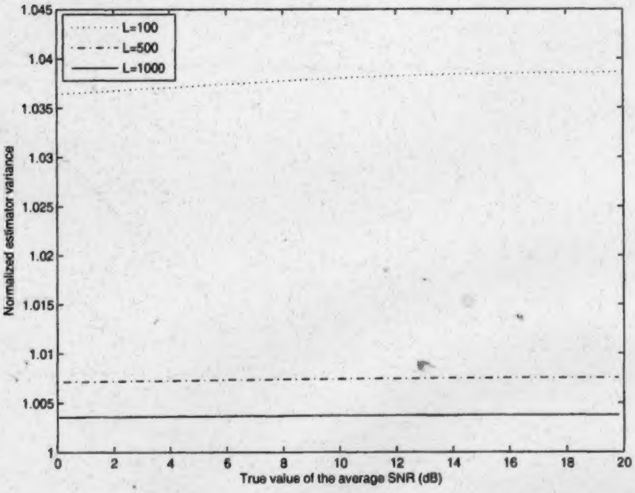

Figure 4.18. The normalized variances (with respect to the corresponding CRLBs) of $\hat{\bar{\rho}}$ for $I = 2$ and $K = 5$ for different sample sizes.

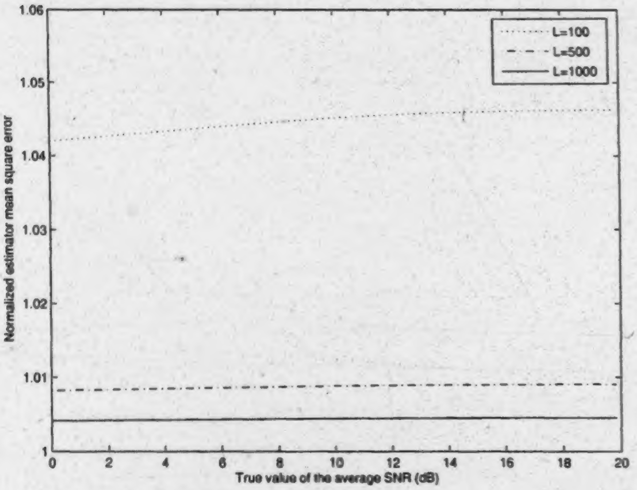

Figure 4.19. The normalized mean square errors (with respect to the corresponding CRLBs) of $\hat{\rho}$ for $I = 2$ and $K = 5$ for different sample sizes.

and the ASNR using correlated samples. In this work, we investigate the problem of estimating the ASNR and the Ricean K parameter jointly in a Ricean fading channel using noisy correlated samples, a case of great practical importance. Two Ricean fading channel models, a time-varying line-of-sight (LOS) component model and a constant LOS component model, are considered [54]. Both data-aided (DA) and non-data-aided (NDA) designs are discussed.

4.4.1 System Model

Consider a system where the data signals experience fast and flat fading. The signal after transmission over the fading channel can be expressed as

$$y(t) = A(t)s(t) + n(t) \tag{4.47}$$

where $A(t)$ is the fading process, $s(t)$ is the transmitted signal, and $n(t)$ is the additive white Gaussian noise (AWGN). In a Ricean fading channel, $A(t)$ is a complex Gaussian random process with mean $E\{A(t)\} = e(t)$ ($e(t) \neq 0$) and autocovariance $E\{[A(t) - e(t)][A(t+\tau) - e(t+\tau)]^*\} =$

$R(t,t+\tau)$. We consider Jake's isotropic scattering model in this paper. Thus, one has [54]

$$R(t,t+\tau) = 2\alpha^2 J_0(2\pi f_D \tau) \tag{4.48}$$

where $2\alpha^2 = E\{|A(t) - e(t)|^2\}$ is the variance of $A(t)$, $J_0(\cdot)$ is the Bessel function of order zero of the first kind [2, p. 358], and f_D is the maximum Doppler shift. The mean, $e(t)$, is actually the LOS component in the Ricean fading channel. One Ricean fading channel model suggests that [54]

$$e(t) = Pe^{j(2\pi \tilde{f}_d t + \Phi_0)} \tag{4.49}$$

where P is the amplitude of the LOS component, $\tilde{f}_d = f_D \cos\theta_0$ is the Doppler shift of the LOS component, θ_0 is the angle of arrival of the LOS component, and Φ_0 is the phase offset of the LOS component. One sees from (4.49) that the LOS component is actually a deterministic sine wave. Since $e(t)$ depends on the time, this model leads to $A(t)$ being a non-stationary fading process and $y(t)$ being a non-stationary received signal. In some applications, a stationary fading process or a stationary received signal may be prefered. In this case, another Ricean fading channel model [54]

$$e(t) = Pe^{j\Phi_0} \tag{4.50}$$

can be used. Comparing (4.50) with (4.49), one sees that (4.50) is actually a special case of (4.49) where $\tilde{f}_d = 0$. Since (4.50) is a constant independent of time, the fading process, and therefore, the received signal are stationary under this model. We will discuss both models in the paper. The transmitted signal can be written as $s(t) = \sum_l b_l p(t - lT)$, where b_l is the l-th transmitted data symbol, $p(t)$ is the shaping pulse with unit symbol energy and T is the symbol interval. The noise in the channel is a complex Gaussian random process with mean zero and autocovariance

$$E\{n(t)n^*(t+\tau)\} = 2\sigma^2\delta(\tau) \tag{4.51}$$

where $\delta(\cdot)$ is the impulse function.

Assume that there is no intersymbol interference. The received signal in (4.47) is matched-filtered and sampled at the time $t = lT$. The l-th sample of the received signal is [55]

$$y_l = A_l b_l + n_l \tag{4.52}$$

where $A_l = A(lT)$ represents the complex Gaussian fading gain in the l-th sample and n_l represents the complex Gaussian noise in the l-th sample. From (4.49) and (4.50), the mean of A_l is

$$e_l = e(lT) = Pe^{j(2\pi \bar{f}_d Tl + \Phi_0)} \tag{4.53}$$

under the assumption of a time-varying LOS component, and is

$$e_l = e(lT) = Pe^{j\Phi_0} \tag{4.54}$$

under the assumption of a constant LOS component. The autocovariance of A_l is

$$E\{(A_l - e_l)(A_{l+h} - e_{l+h})^*\} = 2\alpha^2 J_0(2\pi f_D Th) \tag{4.55}$$

where h is an integer. The mean of n_l is zero, and the autocovariance of n_l is

$$E\{n_l n_{l+h}^*\} = 2\alpha^2 \varepsilon(h) \tag{4.56}$$

where $\varepsilon(h) = 1$ when $h = 0$ and $\varepsilon(h) = 0$ when $h \neq 0$. Note that, in [67]- [72], either A_l $(l = 1, 2, \cdots)$ are assumed to be independent, or y_l $(l = 1, 2, \cdots)$ are assumed to be noiseless, or both, while in this paper and in practice, A_l $(l = 1, 2, \cdots)$ are correlated and y_l $(l = 1, 2, \cdots)$ are noisy, as can be seen from (4.52). Define $K = \frac{P^2}{2\alpha^2}$ as the Ricean K parameter and $\bar{\rho} = 10\log_{10}\left(\frac{P^2+2\alpha^2}{2\sigma^2}\right)$ as ASNR to be estimated. In the sequel, we will propose both DA and NDA estimators for K and $\bar{\rho}$ using L correlated and noisy channel samples.

4.4.2 DA Estimation

In DA estimation, the transmitted data symbols are known. Thus, the received samples in (4.52) can be compensated by the known data symbols to give

$$x_l = A_l + z_l \tag{4.57}$$

where $l = 1, 2, \cdots, L$, $x_l = \frac{y_l}{b_l}$, and $z_l = \frac{n_l}{b_l}$. This is the case when the samples are obtained from the received signals of the modulated or demodulated pilot symbols that may be for synchronizer training or channel estimation [95]. This is also the case when the samples are taken from the received signals of the data symbols and correct data decisions are made with high probability. Assume that $|b_l|^2 = 1$ for $l = 1, 2, \cdots, L$ such that the noise samples in (4.57) are identically distributed. Two Ricean fading channel models are considered as follows.

4.4.2.1 Time-varying LOS Component

In this Ricean fading channel model, the LOS component is time-dependent. From (4.57), together with (4.53), (4.55) and (4.56), one has

$$\mu_1 = E\{x_l\} = Pe^{j(l\omega_d + \Phi_0)} \tag{4.58a}$$

$$\mu_{21} = E\{x_l x_{l+H}^*\} = P^2 e^{-jH\omega_d} + 2\alpha^2 J_0(H\omega_m) \tag{4.58b}$$

$$\mu_{22} = E\{x_l x_{l+2H}^*\} = P^2 e^{-j2H\omega_d} + 2\alpha^2 J_0(2H\omega_m) \tag{4.58c}$$

$$\mu_2 = E\{|x_l|^2\} = P^2 + 2\alpha^2 + 2\sigma^2 \tag{4.58d}$$

where $\omega_m = 2\pi f_D T$ is the normalized maximum Doppler shift in radians, $\omega_d = 2\pi \tilde{f}_d T$ is the normalized Doppler shift of the LOS component in radians, and $H \geq 1$ is an integer. Note that, among the three parameters P, α and σ, the imaginary parts of (4.58b) and (4.58c) depend on P only, the real parts of (4.58b) and (4.58c) depend on both P and α, while (4.58d) depends on all three parameters P, α and σ. This enables us to separate them and to design estimators for each of them. Based on the estimates of P, α and σ, one can then obtain estimators for K and $\bar{\rho}$. Note further that, by properly choosing a large value of the sample lag H such that $J_0(H\omega_m) \approx 0$, nearly independent channel samples used in the previous estimators may still be available from (4.57). However, in doing this, many useful samples will be lost, especially when ω_m is small. Next, we derive moment-based estimators for K and $\bar{\rho}$. We obtain estimators for P^2 and $2\alpha^2$ first.

Denote $\hat{\mu}_1 = \frac{1}{L}\sum_{l=1}^{L} x_l$, $\hat{\mu}_{21} = \frac{1}{L-H}\sum_{l=1}^{L-H} x_l x_{l+H}^*$, $\hat{\mu}_{22} = \frac{1}{L-2H}\sum_{l=1}^{L-2H} x_l x_{l+2H}^*$ and $\hat{\mu}_2 = \frac{1}{L}\sum_{l=1}^{L} |x_l|^2$ as estimates of μ_1, μ_{21}, μ_{22} and μ_2, respectively. From (4.58b) and (4.58c), one moment-based estimator for ω_d, $\hat{\omega}_d$, is

$$\hat{\omega}_d = \frac{1}{H} \arccos \frac{Im\{\hat{\mu}_{22}\}}{2Im\{\hat{\mu}_{21}\}} \tag{4.59}$$

where the imaginary part of the autocorrelation function of the Ricean fading process is used. Based on (4.59), the mean power of the LOS component, P^2, can be estimated using the imaginary part of (4.58b) as

$$\hat{P}^2 = \left| \frac{Im\{\hat{\mu}_{21}\}}{\sin(H\hat{\omega}_d)} \right| \tag{4.60}$$

and the mean power of the scattering components, $2\alpha^2$, can be estimated using the real part of (4.58b) as

$$2\hat{\alpha}^2 = \left| \frac{Re\{\hat{\mu}_{21}\} - \hat{P}^2\cos(H\hat{\omega}_d)}{J_0(H\omega_m)} \right| \tag{4.61}$$

where the magnitude is used, recognizing the fact that both P^2 and $2\alpha^2$ are positive numbers. In (4.61), ω_m is assumed to be known. Finally, the DA moment-based estimators for K and $\bar{\rho}$ can be derived by using the estimates of P^2 and $2\alpha^2$ to give, respectively,

$$\hat{K}_{DA-LOSA} = \frac{\hat{P}^2}{2\hat{\alpha}^2} \tag{4.62}$$

and

$$\hat{\bar{\rho}}_{DA-LOSA} = 10\log_{10}\left[\frac{\hat{P}^2 + 2\hat{\alpha}^2}{\hat{\mu}_2 - \hat{P}^2 - 2\hat{\alpha}^2} \right] \tag{4.63}$$

where $\hat{P}^2$ is given by (4.60), $2\hat{\alpha}^2$ is given by (4.61), and (4.58d) is used. One sees that the estimators in (4.62) and (4.63) are derived by using the second order statistics of the received signals only, as the autocorrelation function of the Ricean fading process contains enough information for estimation. Specifically, the moment-based estimator for K in (4.62) uses μ_{21} and μ_{22}, and the moment-based estimator for $\bar{\rho}$ in (4.63) uses μ_{21}, μ_{22} and μ_2. One also sees that the estimators in (4.62) and (4.63) actually represent a class of moment-based estimators, as the values of the sample lags, H and $2H$, can be any positive integers. Estimators with different values of H and $2H$ will have different performances, as will be discussed later. In our derivation, the sample lags of H and $2H$ are used to facilitate the calculation of $\hat{\omega}_d$ in (4.59). Several comments are made as follows.

First, in (4.61), ω_m is assumed available. It is well known that $\tilde{f}_m = \frac{v}{\lambda}$, where v is the mobile speed and λ is the wavelength [54]. Since $\omega_m = 2\pi\tilde{f}_m T$, given a specific communication system and a specific symbol rate, ω_m is solely determined by the mobile speed. Estimation of the mobile speed has been well studied by several researchers [96]- [99]. In [96] and [97], mobile velocity estimators that are independent of, or robust with respect to, the Ricean K parameter have been derived. In [98] and [99], the authors proposed mobile velocity estimators for a Rayleigh fading channel, which can be easily extended to a Ricean fading channel. All these estimators have good performances. As reported in [99], the standard deviation of the estimate is within 2.4 Km/h. Therefore, the value of ω_m can be accurately estimated. Moreover, it will be shown later that by properly choosing the value of H, the performances of the derived estimators are not sensitive to errors in the estimation of ω_m.

Second, in [9], the authors have derived another moment-based estimator for ω_d as

$$\hat{\omega}_d = arg \max_{\omega} \left\{ \left| \frac{1}{L} \sum_{l=1}^{L} x_l e^{-jl\omega} \right| \right\} \tag{4.64}$$

and another moment-based estimator for P^2 as

$$\hat{P}^2 = \left| \frac{1}{L} \sum_{l=1}^{L} x_l e^{-jl\hat{\omega}_d} \right|^2 \tag{4.65}$$

for a noiseless Ricean fading channel. These estimators could also be used in the noisy case, as they use the first order moment of the sample and the mean of the noise is zero. However, one sees from (4.64) that this estimator requires a search over all possible values of w_d, which has a complexity of $L \log_2 L$ for each w. Our proposed estimator in (4.59) saves estimation time as well as estimator complexity. Also, denoting $\Delta = \omega_d - \hat{\omega}_d$ as the estimation error in the estimation of ω_d, the estimator in (4.65) is very sensitive to Δ. From (4.65), one has $\hat{P}^2 \approx P^2 \left| \frac{\sin(\frac{L+1}{2}\Delta)}{L \sin(\frac{\Delta}{2})} \right|^2$ when the ratio of P^2 to $2\sigma^2 + 2\alpha^2$ is large. At $L = 1000$, $\Delta = 0.01$ will give an estimate of $\hat{P}^2 = 0.037P^2$, which is far from the true value of P^2. Thus, an estimate of w_d with very high accuracy is required by (4.65). On the contrary, our proposed estimator in (4.60) is robust to Δ, giving higher quality estimates.

Finally, it can be seen from [59] that $2\sigma^2$ and $2\alpha^2$ always appear together as $2\sigma^2 + 2\alpha^2$ in moments of the sample envelope on a noisy Ricean fading channel. As a result, one cannot use the sample envelope to estimate $\bar{\rho}$ and K, as one did in the noiseless case. We use the autocorrelation function of the samples in this paper. Since the samples from the Ricean fading process are correlated while the samples from the noise process are uncorrelated, this is probably the only way of differentiating $2\sigma^2$ and $2\alpha^2$, without the use of an independent noise power estimator. Next, we discuss moment-based estimators for the case of a constant LOS component.

4.4.2.2 Constant LOS Component

In this Ricean fading channel model, the LOS component is constant, giving a stationary Ricean fading process. The moments of the received samples are obtained from (4.57) as

$$\mu_1 = E\{x_l\} = Pe^{j\Phi_0} \tag{4.66a}$$

$$\mu_{21} = E\{x_l x_{l+H}^*\} = P^2 + 2\alpha^2 J_0(H\omega_m) \tag{4.66b}$$

$$\mu_2 = E\{|x_l|^2\} = P^2 + 2\alpha^2 + 2\sigma^2 \tag{4.66c}$$

where the notations are defined as before. Similarly, one sees that, among the three parameters P, α and σ, (4.66a) depends on P only, (4.66b) depends on both P and α, and (4.66c) depends on all three parameters, enabling us to estimate P, α and σ separately. One also sees that μ_1 in (4.66a) doesn't depend on the sample index l and the imaginary part of the autocorrelation function in (4.66b) disappears, as a result of the stationarity of the Ricean fading process. Similar to before, we derive estimators for P^2 and $2\alpha^2$ first.

Using (4.66a) and (4.66b), a moment-based estimator for P^2, $\hat{P^2}$, is

$$\hat{P^2} = |\hat{\mu}_1|^2 \tag{4.67}$$

and a moment-based estimator for $2\alpha^2$, $2\hat{\alpha^2}$, is

$$2\hat{\alpha^2} = \left| \frac{Re\{\hat{\mu}_{21}\} - \hat{P^2}}{J_0(H\omega_m)} \right| \tag{4.68}$$

where $\hat{\mu}_1$ and $\hat{\mu}_{21}$ are defined as before. The moment-based estimators for the Ricean K parameter and the local average SNR can be obtained, respectively, as

$$\hat{K}_{DA-LOSB} = \frac{\hat{P^2}}{2\hat{\alpha^2}} \tag{4.69}$$

and

$$\hat{\gamma}_{DA-LOSB} = 10\log_{10}\left[\frac{\hat{P}^2 + 2\hat{\alpha}^2}{\hat{\mu}_2 - \hat{P}^2 - 2\hat{\alpha}^2}\right] \tag{4.70}$$

where $\hat{P}^2$ is given by (4.67), $2\hat{\alpha}^2$ is given by (4.68) and $\hat{\mu}_2$ is defined as before. As in the previous subsection, the value of ω_m in (4.68) is assumed known. Note that estimation of the Ricean K parameter and the local average SNR under the assumption of constant LOS component is much simpler than under the assumption of time-varying LOS component, as one doesn't need to estimate the Doppler shift of the LOS component in this case. Note further that the moment-based estimator for K in (4.69) uses μ_1 and μ_{21}, and the moment-based estimator for $\bar{\rho}$ in (4.70) uses μ_1, μ_{21} and μ_2. Next, we derive NDA estimators for K and $\bar{\rho}$. To the best of the authors' knowledge, no NDA estimators for the Ricean K parameter have been described in the literature, nor have NDA estimators designed for noisy fading channels been examined.

4.4.3 NDA Estimation

In NDA estimation, the transmitted data symbols are unknown. As a result, the received samples cannot be compensated by the data symbols and the properties of the data sequence, b_l $(l = 1, 2, \cdots)$, will affect the blind estimation of the Ricean K parameter and the local average SNR. As done in [1, eq. (4.4-5)], we assume that the data sequence, b_l $(l = 1, 2, \cdots)$, is wide-sense stationary with mean d and autocorrelation function

$$\Psi(h) = E\{b_l b_{l+h}^*\}. \tag{4.71}$$

As will be seen later, our NDA estimators require that $\Psi(h)$ be non-zero at $h = H$ and $h = 2H$, where H and $2H$ are the sample lags used in the estimation.

It is well known that the autocorrelation function of a wide-sense stationary process is the sum of the squared mean of the process and the autocovariance function of the process [34, eq. (9-9)]. In a communication system where coding is introduced, if the received signals are sampled before decoding, the autocovariance function of the data sequence in these samples will be non-zero, which in turn makes the autocorrelation function of the data sequence non-zero. Also, in a communication system where modulation with memory is used, if the received signals are sampled before demodulation, the autocovariance function of the data sequence in these samples will be non-zero, giving a non-zero autocorrelation function as well [1]. Even if no coding or modulation with memory is adopted in the communication system and the autocovariance function of the data sequence is zero, by using signaling schemes such as on-off keying and frequency-shift-keying, the mean of the data symbols may still be non-zero, leading to a non-zero autocorrelation function of the data sequence as well. In fact, the autocorrelation of the data sequence is often used in practical systems to control the power spectrum of the modulated signals [1]. With a zero autocorrelation function of the data sequence, the power spectrum of the modulated signals will also be zero, an impractical case in realistic systems. We assume that $\Psi(h)$ is non-zero and is known at $h = H$ and $h = 2H$ in the following derivation.

4.4.3.1 Time-varying LOS Component

Again, we begin with the Ricean fading channel model where the LOS component is time-varying. Assuming that the data sequence is independent of the fading process, from (4.52), we have

$$v_1 = E\{y_l\} = Pe^{j(l\omega_d+\Phi_0)}d \tag{4.72a}$$

$$v_{21} = E\{y_l y_{l+H}^*\} = [P^2 e^{-jH\omega_d} + 2\alpha^2 J_0(H\omega_m)]\Psi(H) \tag{4.72b}$$

$$v_{22} = E\{y_l y_{l+2H}^*\} = [P^2 e^{-j2H\omega_d} + 2\alpha^2 J_0(2H\omega_m)]\Psi(2H) \tag{4.72c}$$

$$v_2 = E\{|y_l|^2\} = (P^2 + 2\alpha^2)\Psi(0) + 2\sigma^2 \tag{4.72d}$$

where d is the mean of the data sequence, $\Psi(H)$, $\Psi(2H)$ and $\Psi(0)$ are the values of $\Psi(h)$ at $h = H$, $h = 2H$ and $h = 0$, respectively, and ω_m and ω_d are defined as before. One sees that the only difference between the second order moments of the DA samples in (4.58b)-(4.58d) and the second order moments of the NDA samples in (4.72b)-(4.72d) is that (4.72b)-(4.72d) depend on the autocorrelation function of the data sequence, as the data symbols are unknown in this case. Denote $\hat{v}_1 = \frac{1}{L}\sum_{l=1}^{L} y_l$, $\hat{v}_{21} = \frac{1}{L-H}\sum_{l=1}^{L-H} y_l y_{l+H}^*$, $\hat{v}_{22} = \frac{1}{L-2H}\sum_{l=1}^{L-2H} y_l y_{l+2H}^*$ and $\hat{v}_2 = \frac{1}{L}\sum_{l=1}^{L} |y_l|^2$ as estimates of v_1, v_{21}, v_{22} and v_2, respectively. Using similar techniques as previously, we can derive the moment-based estimator for ω_d, $\hat{\omega}_d$, as

$$\hat{\omega}_d = \frac{1}{H}\arccos\left\{\frac{\Psi(H)Im\{\hat{v}_{22}\}}{2\Psi(2H)Im\{\hat{v}_{21}\}}\right\}. \tag{4.73}$$

The moment-based estimators for P^2 and $2\alpha^2$, $\hat{P^2}$ and $2\hat{\alpha}^2$, can be obtained by using (4.73), together with (4.72b), as

$$\hat{P^2} = \left|\frac{Im\{\hat{v}_{21}\}}{\Psi(H)\sin(H\hat{\omega}_d)}\right| \tag{4.74}$$

and

$$2\hat{\alpha}^2 = \left|\frac{Re\{\hat{v}_{21}\} - \hat{P^2}\cos(H\hat{\omega}_d)\Psi(H)}{\Psi(H)J_0(H\omega_m)}\right|. \tag{4.75}$$

Finally, the NDA moment-based estimators for K and $\bar{\rho}$ are

$$\hat{K}_{NDA-LOSA} = \frac{\hat{P^2}}{2\hat{\alpha}^2} \tag{4.76}$$

and

$$\hat{\gamma}_{NDA-LOSA} = 10\log_{10}\left[\frac{\hat{P^2} + 2\hat{\alpha}^2}{\hat{v}_2 - \Psi(0)(\hat{P^2} + 2\hat{\alpha}^2)}\right] \tag{4.77}$$

where $\hat{P^2}$ is given by (4.74) and $2\hat{\alpha}^2$ is now given by (4.75). Note that, again, the estimators in (4.76) and (4.77) use the second order statistics of the received samples only, as the autocorrelation function contains enough information about the unknown parameters.

4.4.3.2 Constant LOS Component

In this Ricean fading channel model, the LOS component is constant. From (4.52), one has

$$v_1 = E\{y_l\} = Pe^{j\Phi_0}d \tag{4.78a}$$

$$v_{21} = E\{y_l y_{l+H}^*\} = [P^2 + 2\alpha^2 J_0(H\omega_m)]\Psi(H) \tag{4.78b}$$

$$v_{22} = E\{y_l y_{l+2H}^*\} = [P^2 + 2\alpha^2 J_0(2H\omega_m)]\Psi(2H) \tag{4.78c}$$

$$v_2 = E\{|y_l|^2\} = [P^2 + 2\alpha^2]\Psi(0) + 2\sigma^2 \tag{4.78d}$$

where, again, d is the mean of the data sequence. If $d \neq 0$, the moment-based estimators for P^2 and $2\alpha^2$ are derived from (4.78a) and (4.78b) as

$$\hat{P^2} = \frac{|\hat{v}_1|^2}{d^2} \tag{4.79}$$

and

$$2\hat{\alpha}^2 = \left| \frac{Re\{\hat{v}_{21}\} - \Psi(H)\hat{P}^2}{\Psi(H)J_0(H\omega_m)} \right|. \tag{4.80}$$

Using (4.79) and (4.80), the moment-based estimator for K, $\hat{K}_{NDA-LOSB}$, is

$$\hat{K}_{NDA-LOSB} = \frac{\hat{P}^2}{2\hat{\alpha}^2} \tag{4.81}$$

and the moment-based estimator for $\bar{\rho}$, $\hat{\bar{\rho}}_{NDA-LOSB}$, is

$$\hat{\bar{\rho}}_{NDA-LOSB} = 10\log_{10}\left[\frac{\hat{P}^2 + 2\hat{\alpha}^2}{\hat{v}_2 - \Psi(0)(\hat{P}^2 + 2\hat{\alpha}^2)}\right] \tag{4.82}$$

where $\hat{P}^2$ is given by (4.79) and $2\hat{\alpha}^2$ is given by (4.80). Note that $\hat{K}_{NDA-LOSB}$ uses v_1 and v_{21}, while $\hat{\bar{\rho}}_{NDA-LOSB}$ uses v_1, v_{21} and v_2. Both estimators fail in the case when $d = 0$, as the first order moment of the received samples will become zero and will not be usable in the estimation. We propose estimators that can be used for $d = 0$ in the following. From (4.78b) and (4.78c), one has moment-based estimators for P^2 and $2\alpha^2$, $\hat{P}^2$ and $2\hat{\alpha}^2$, as

$$\hat{P}^2 = \left| \frac{J_0(H\omega_m)}{J_0(H\omega_m) - J_0(2H\omega_m)} \cdot \frac{Re\{\hat{v}_{22}\}}{\Psi(2H)} - \frac{J_0(2H\omega_m)}{J_0(H\omega_m) - J_0(2H\omega_m)} \cdot \frac{Re\{\hat{v}_{21}\}}{\Psi(H)} \right| \tag{4.83}$$

and

$$2\hat{\alpha}^2 = \left| \frac{\Psi(2H)Re\{\hat{v}_{21}\} - \Psi(H)Re\{\hat{v}_{22}\}}{\Psi(H)\Psi(2H)[J_0(H\omega_m) - J_0(2H\omega_m)]} \right|. \tag{4.84}$$

Using (4.83) and (4.84), the moment-based estimator for K, $\hat{K}_{NDA-LOSB2}$, is

$$\hat{K}_{NDA-LOSB2} = \frac{\hat{P}^2}{2\hat{\alpha}^2} \tag{4.85}$$

and the moment-based estimator for $\bar{\rho}$, $\hat{\bar{\rho}}_{NDA-LOSB2}$, is

$$\hat{\bar{\rho}}_{NDA-LOSB2} = 10\log_{10}\left[\frac{\hat{P}^2 + 2\hat{\alpha}^2}{\hat{v}_2 - \Psi(0)(\hat{P}^2 + 2\hat{\alpha}^2)}\right] \tag{4.86}$$

where $\hat{P}^2$ and $2\hat{\alpha}^2$ are given by (4.83) and (4.84), respectively. Note that the estimators in (4.85) and (4.86) apply to the case when $d \neq 0$ as well as the case when $d = 0$, since they don't use the first order moment of the received samples.

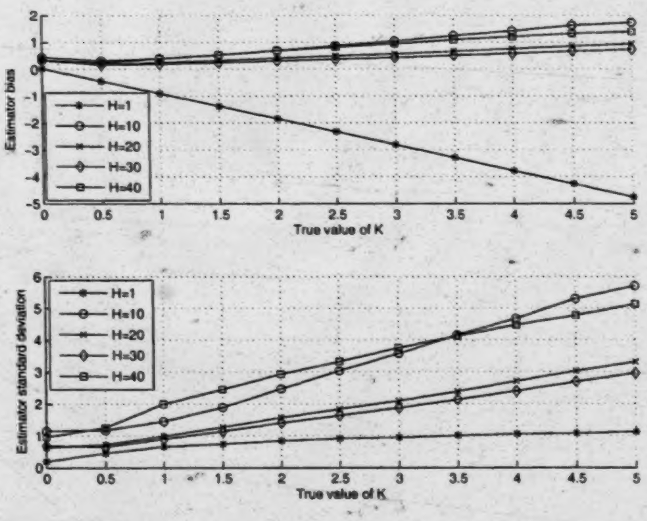

Figure 4.20. The performance of $\hat{K}_{DA-LOSA}$ for different values of the sample lag with $L = 1024$ and $\bar{\rho} = 10\, dB$.

4.4.4 Numerical Results

Here, the performances of the estimators for K and $\bar{\rho}$ are examined and compared with the estimator in (3.39) and the estimator in (4.41), respectively. For convenience, we denote the estimator in (3.39) as $\hat{K}_{old}$ and the estimator in (4.41) as $\hat{\bar{\rho}}_{old}$. We study a communication system with a carrier frequency of 2 GHz. The mobile travels at a speed of 100 Km/h and the received signal is sampled at a symbol rate of 25 kb/s. Without loss of generality, let $\Phi_0 = \frac{\pi}{4}$ and $\theta_0 = \frac{\pi}{6}$ be the phase offset and the angle of arrival of the LOS component, respectively. The total mean power of the fading is fixed to 1.0, while K varies from 0 to 5 with a step size of 0.5. The true value of $\bar{\rho}$ is set from 0 dB to 15 dB with increments of 1 dB. In the NDA estimation, binary frequency shift keying is assumed to illustrate the effect of unknown data symbols on the estimator performance, where $b_l = 0$ and $b_l = 1$ are sent with equal probabilities and b_l ($l = 1, 2, \cdots, L$) are mutually uncorrelated. Other schemes can be examined accordingly.

We discuss the effect of H on the estimator performance first. We use $\hat{K}_{DA-LOSA}$ as an example.

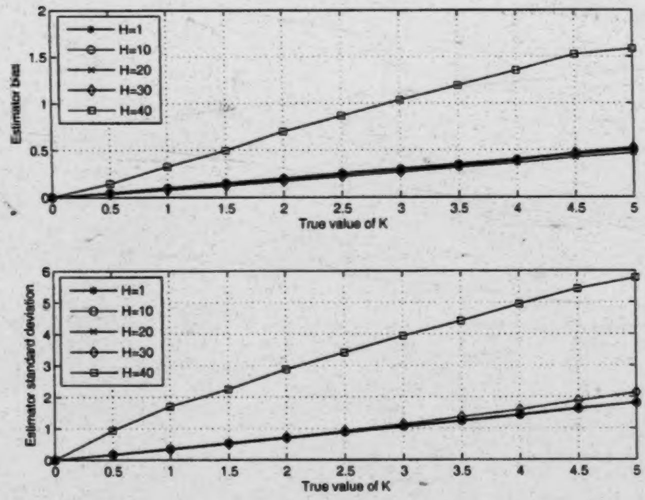

Figure 4.21. The performance of $\hat{K}_{DA-LOSB}$ for different values of the sample lag with $L = 1024$ and $\bar{\rho} = 10\ dB$.

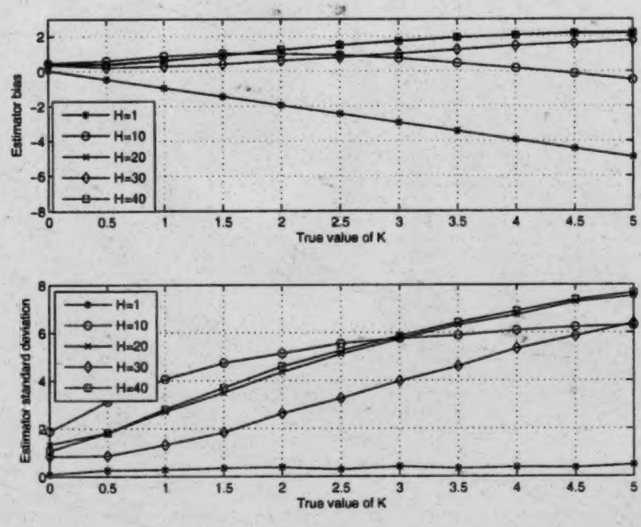

Figure 4.22. The performance of $\hat{K}_{NDA-LOSA}$ for different values of the sample lag with $L = 1024$ and $\bar{\rho} = 10\ dB$.

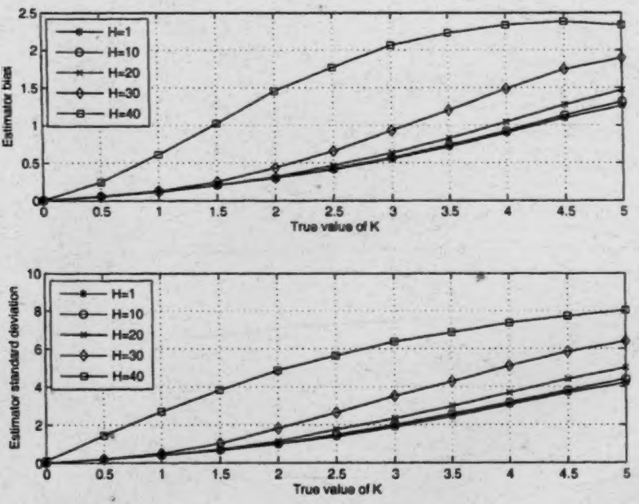

Figure 4.23. The performance of $\hat{K}_{NDA-LOSB}$ for different values of the sample lag with $L = 1024$ and $\bar{\rho} = 10\ dB$.

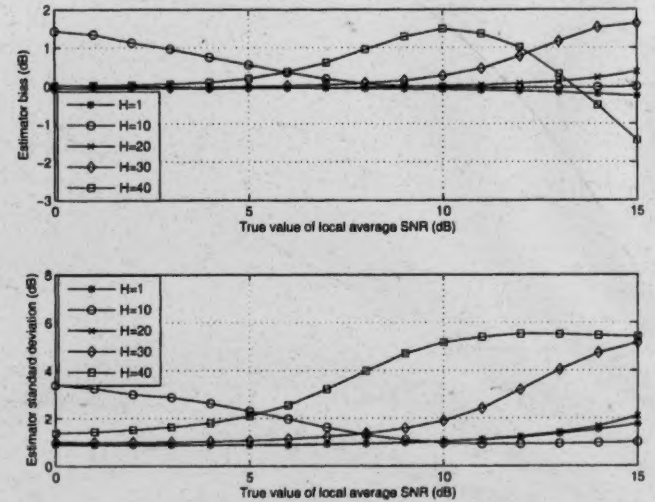

Figure 4.24. The performance of $\hat{\bar{\rho}}_{DA-LOSA}$ for different values of the sample lag with $L = 1024$ and $K = 5$.

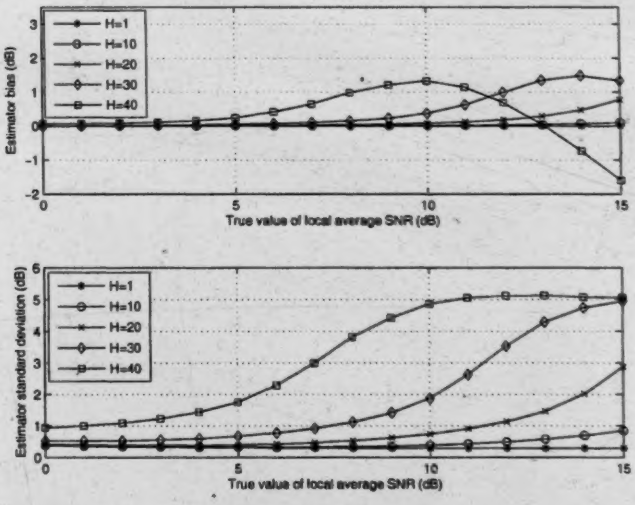

Figure 4.25. The performance of $\hat{\hat{\rho}}_{DA-LOSB}$ for different values of the sample lag with $L = 1024$ and $K = 5$.

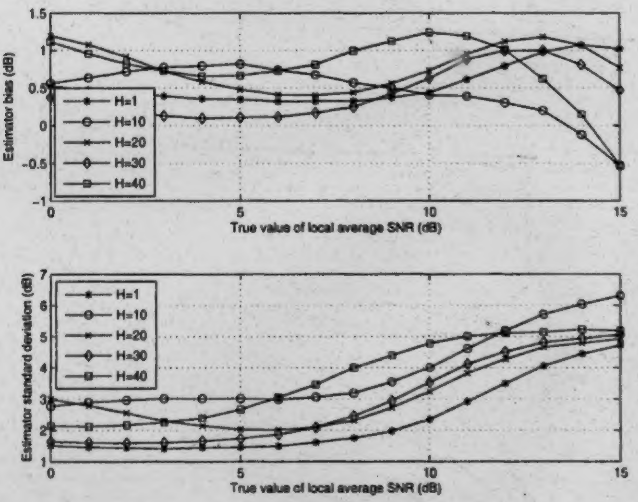

Figure 4.26. The performance of $\hat{\hat{\rho}}_{NDA-LOSA}$ for different values of the sample lag with $L = 1024$ and $K = 5$.

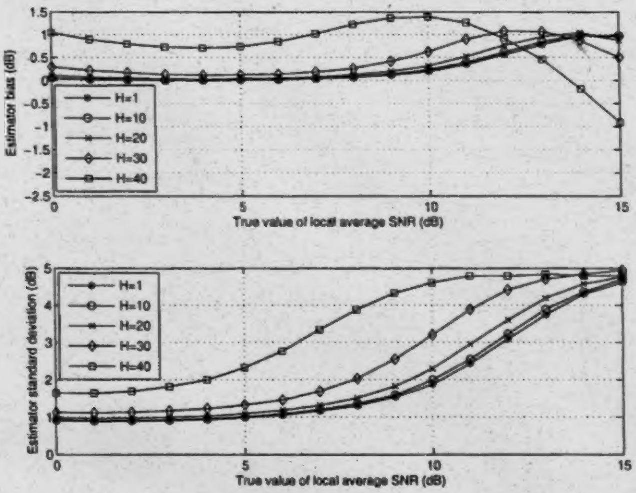

Figure 4.27. The performance of $\hat{\mathring{\rho}}_{NDA-LOSB}$ for different values of the sample lag with $L = 1024$ and $K = 5$.

The value of H affects the simulated performance of $\hat{K}_{DA-LOSA}$ in three ways. First, it determines the value of $H\omega_d$. For a fixed ω_d, the imaginary parts of μ_{21} and μ_{22} will be close to zero and the estimators (4.59) and (4.60) will be vulnerable to noise when H is small. Second, it determines the value of $H\omega_m$. For a fixed ω_m, the Bessel functions in μ_{21} and μ_{22} will be close to zero and estimator (4.61) will be vulnerable to noise when H is large. Third, it determines the effective sample sizes of $L-H$ and $L-2H$. A larger H corresponds to smaller effective sample sizes. The overall performance of the estimator depends on all the three factors. Fig. 4.20 shows the effect of H on the performance of $\hat{K}_{DA-LOSA}$. Five values of the sample lag, $H = 1$, $H = 10$, $H = 20$, $H = 30$ and $H = 40$ are examined. One sees that $H = 30$ gives the best overall performance of $\hat{K}_{DA-LOSA}$ among all values of the sample lag evaluated. Similarly, one can see from Figs. 4.21 to 4.27 that $H = 30$ gives the best overall performance of $\hat{K}_{NDA-LOSA}$, $H = 20$ gives the best overall performance of $\hat{K}_{DA-LOSB}$, and $H = 1$ gives the best overall performance of $\hat{K}_{NDA-LOSB}$, $\hat{\mathring{\rho}}_{DA-LOSA}$, $\hat{\mathring{\rho}}_{DA-LOSB}$, $\hat{\mathring{\rho}}_{NDA-LOSA}$ and $\hat{\mathring{\rho}}_{NDA-LOSB}$, among the values of sample lag examined.

Next, we discuss the effect of the estimation accuracy of ω_m on the estimator performance.

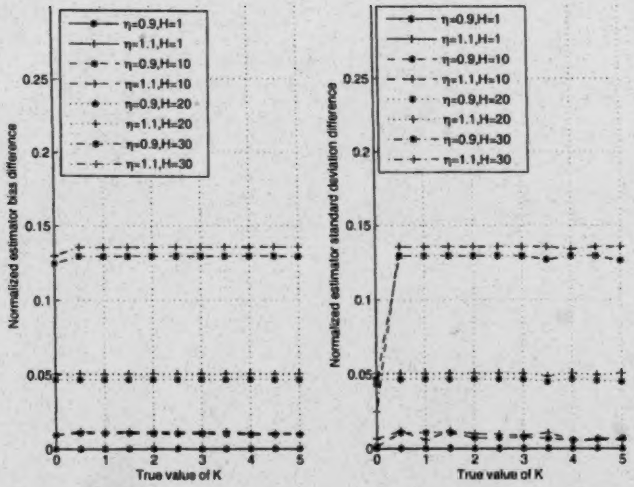

Figure 4.28. The performance of $\hat{K}_{DA-LOSA}$ for mobile speed estimates having different accuracies at $L = 1024$, $\bar{\rho} = 10\,dB$.

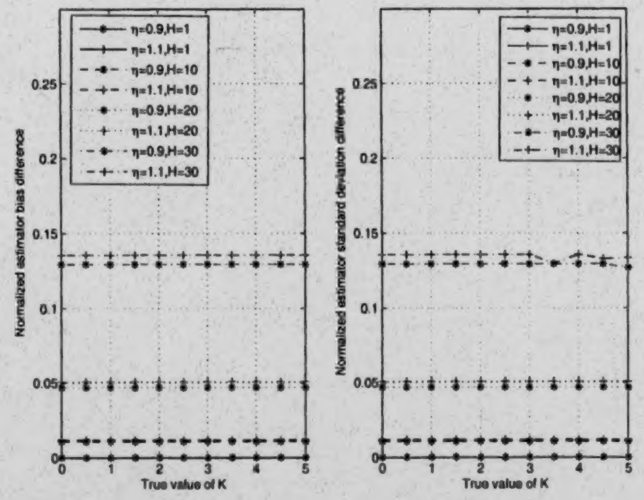

Figure 4.29. The performance of $\hat{K}_{DA-LOSB}$ for mobile speed estimates having different accuracies at $L = 1024$ and $\bar{\rho} = 10\,dB$.

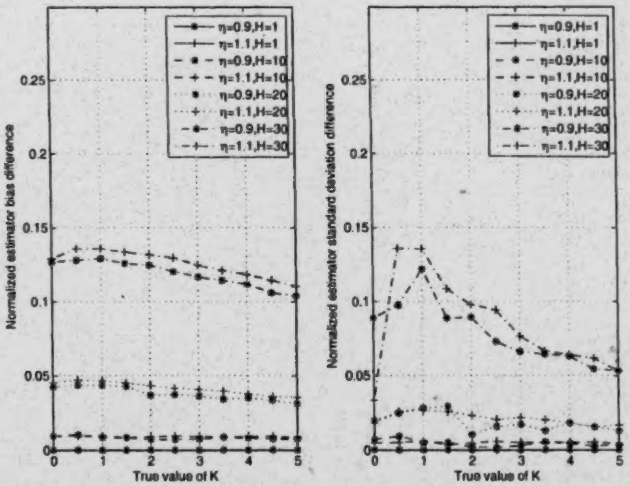

Figure 4.30. The performance of $\hat{K}_{NDA-LOSA}$ for mobile speed estimates having different accuracies at $L = 1024$ and $\bar{\rho} = 10\,dB$.

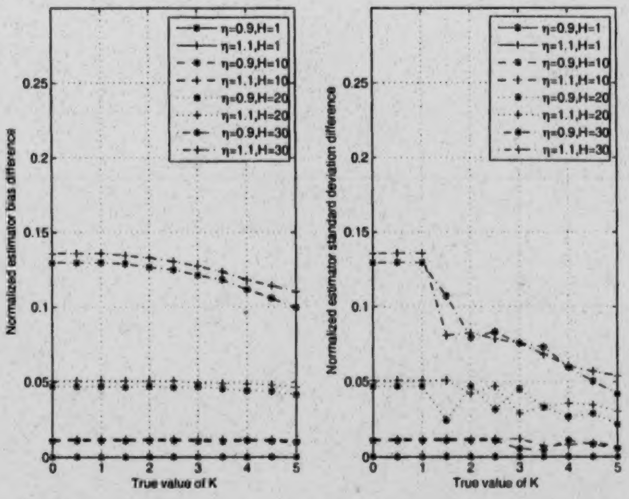

Figure 4.31. The performance of $\hat{K}_{NDA-LOSB}$ for mobile speed estimates having different accuracies at $L = 1024$ and $\bar{\rho} = 10\,dB$.

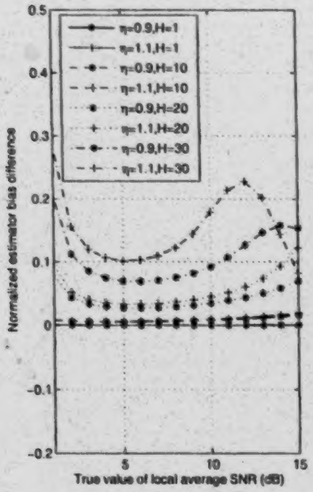
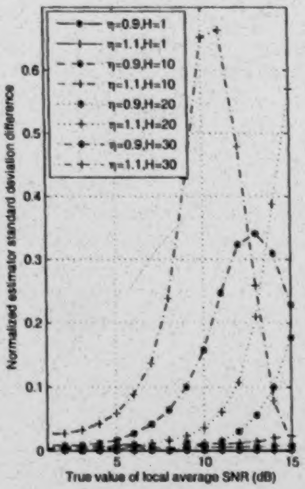

Figure 4.32. The performance of $\hat{\rho}_{DA-LOSA}$ for mobile speed estimates having different accuracies at $L = 1024$ and $K = 5$.

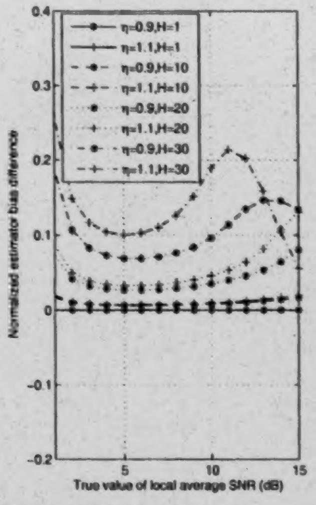
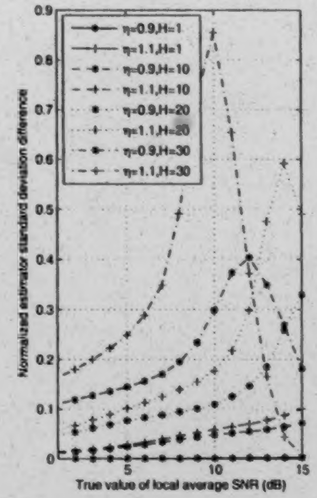

Figure 4.33. The performance of $\hat{\rho}_{DA-LOSB}$ for mobile speed estimates having different accuracies at $L = 1024$ and $K = 5$.

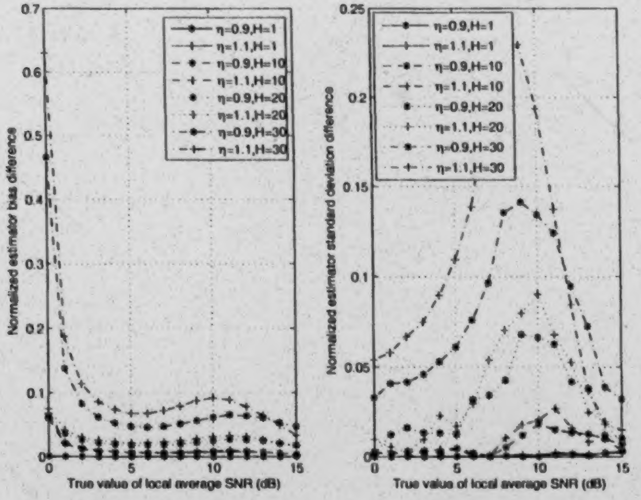

Figure 4.34. The performance of $\hat{\bar{\rho}}_{NDA-LOSA}$ for mobile speed estimates having different accuracies at $L = 1024$ and $K = 5$.

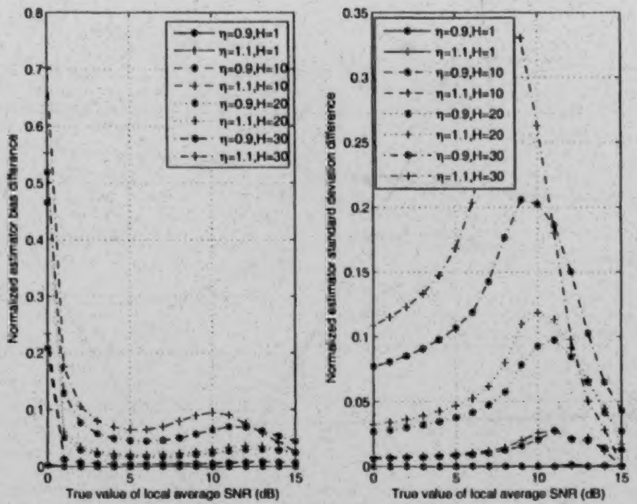

Figure 4.35. The performance of $\hat{\bar{\rho}}_{NDA-LOSB}$ for mobile speed estimates having different accuracies at $L = 1024$ and $K = 5$.

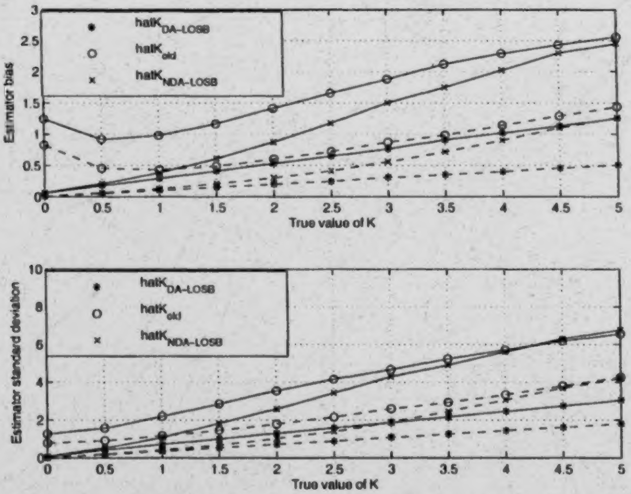

Figure 4.36. Performances of $\hat{K}_{DA-LOSB}$, $\hat{K}_{old}$ and $\hat{K}_{NDA-LOSB}$ with $L = 1024$ (dashed line) and $L = 512$ (solid line) at $\bar{\rho} = 10\,dB$, assuming a constant LOS fading component.

Again, we use $\hat{K}_{DA-LOSA}$ as an example. Fig. 4.28 shows the effect of the estimation accuracy of ω_m on the performance of $\hat{K}_{DA-LOSA}$, where η is defined as the ratio of the mobile speed estimate to its true value and the normalization is taken with respect to the performance of $\hat{K}_{DA-LOSA}$ when $\eta = 1$. One sees that $\hat{K}_{DA-LOSA}$ is more sensitive to $\eta = 1.1$ than to $\eta = 0.9$. One also sees that the effect of the estimation accuracy of ω_m depends on H. The smaller H is, the less sensitive the estimator performance will be to the mobile speed estimation error. From Figs. 4.29 to 4.35, one sees that this observation is valid for other estimators as well. As a compromise between robustness and optimality, in the following simulation, we use $H = 20$ for $\hat{K}_{DA-LOSA}$ and $\hat{K}_{NDA-LOSA}$, and $H = 1$ for other estimators.

Figs. 4.36 and 4.37 show the performances of $\hat{K}_{DA-LOSB}$ and $\hat{K}_{NDA-LOSB}$ for different values of $\bar{\rho}$ and N. One sees that $\hat{K}_{DA-LOSB}$ has a bias between 0.1 and 1.3 and a standard deviation between 0.0 and 3.0 when $L = 512$ and $\bar{\rho} = 10\,dB$, a bias between 0.0 and 0.5 and a standard deviation between 0.0 and 1.9 when $L = 1024$ and $\bar{\rho} = 10\,dB$, a bias between 0.1 and 1.4 and a standard

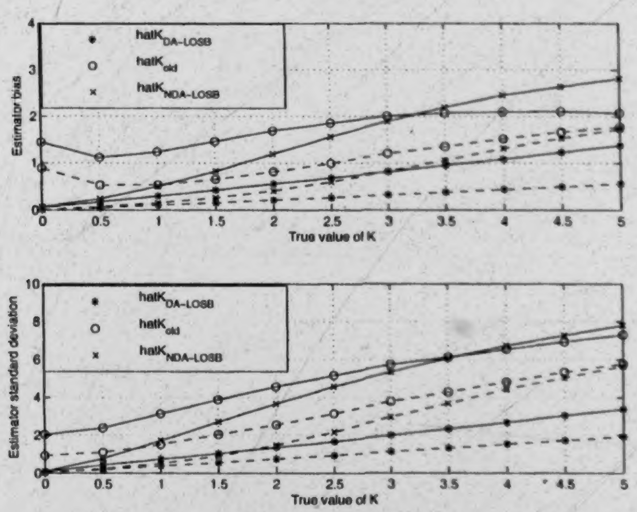

Figure 4.37. Performances of $\hat{K}_{DA-LOSB}$, $\hat{K}_{old}$ and $\hat{K}_{NDA-LOSB}$ with $L = 1024$ (dashed line) and $L = 512$ (solid line) at $\bar{\rho} = 5\ dB$, assuming a constant LOS fading component.

deviation between 0.0 and 3.3 when $L = 512$ and $\bar{\rho} = 5\ dB$, and a bias between 0.0 and 0.6 and a standard deviation between 0.0 and 2.0 when $L = 1024$ and $\bar{\rho} = 5\ dB$, for all the values of K considered. Hence, $\hat{K}_{DA-LOSB}$ has good performance, especially when the true value of K is small. The estimator performance improves as the sample size and/or the local average SNR increases. Comparing $\hat{K}_{DA-LOSB}$ with $\hat{K}_{NDA-LOSB}$, one observes that $\hat{K}_{DA-LOSB}$ outperforms $\hat{K}_{NDA-LOSB}$ for all the cases discussed. Also, comparing $\hat{K}_{DA-LOSB}$ with $\hat{K}_{old}$, one sees that $\hat{K}_{DA-LOSB}$ outperforms $\hat{K}_{old}$ greatly, as expected.

Figs. 4.38 and 4.39 show the performances of $\hat{K}_{DA-LOSA}$ and $\hat{K}_{NDA-LOSA}$ at different values of $\bar{\rho}$ and L. Similar observations can be made. Again, the performance of $\hat{K}_{DA-LOSA}$ improves as L or(and) $\bar{\rho}$ increase(s). Comparing $\hat{K}_{DA-LOSA}$ with $\hat{K}_{old}$, one sees that our correlated-sample-based estimator $\hat{K}_{DA-LOSA}$ outperforms the independent-sample-based estimator $\hat{K}_{old}$ in [59] for all cases considered. Also, comparing $\hat{K}_{DA-LOSA}$ with $\hat{K}_{NDA-LOSA}$, one sees that $\hat{K}_{DA-LOSA}$ performs much better than $\hat{K}_{NDA-LOSA}$ for most cases. However, when $\bar{\rho} = 5\ dB$ and $L = 512$, $\hat{K}_{DA-LOSA}$

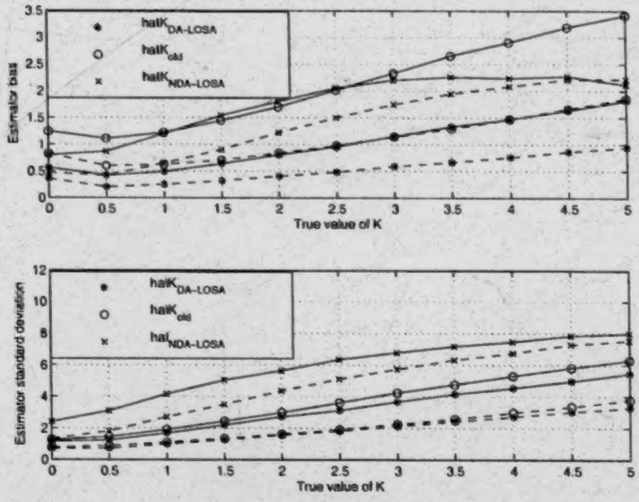

Figure 4.38. Performances of $\hat{K}_{DA-LOSA}$, $\hat{K}_{old}$ and $\hat{K}_{NDA-LOSA}$ with $L = 1024$ (dashed line) and $L = 512$ (solid line) at $\bar{\rho} = 10\,dB$, assuming a time-varying LOS fading component.

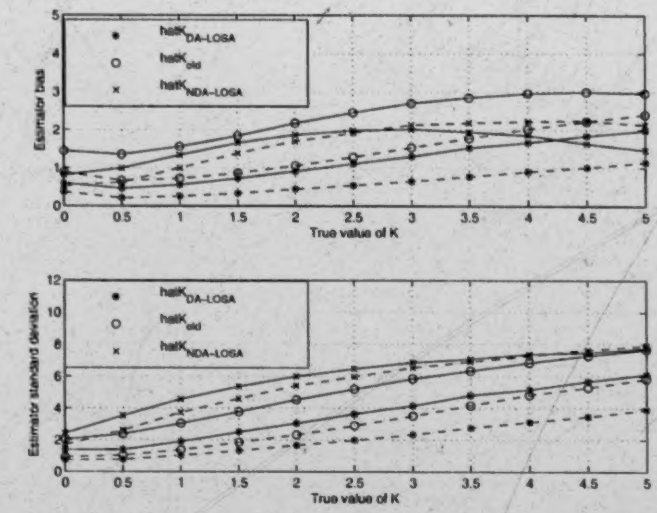

Figure 4.39. Performances of $\hat{K}_{DA-LOSA}$, $\hat{K}_{old}$ and $\hat{K}_{NDA-LOSA}$ with $L = 1024$ (dashed line) and $L = 512$ (solid line) at $\bar{\rho} = 5\,dB$, assuming a time-varying LOS fading component.

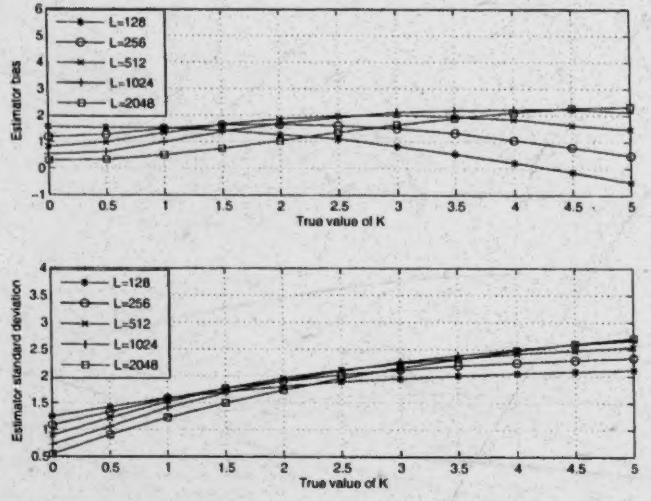

Figure 4.40. Performances of $\hat{K}_{NDA-LOSA}$ at $\bar{\rho} = 5\ dB$ for different sample sizes, assuming a time-varying LOS fading component.

has a smaller standard deviation but a larger bias than $\hat{K}_{NDA-LOSA}$ for $4.2 \leq K \leq 5.0$. Also, when $4.2 \leq K \leq 5.0$, $\hat{K}_{NDA-LOSA}$ with $L = 512$ has a smaller bias as well as a smaller standard deviation than $\hat{K}_{NDA-LOSA}$ with $N = 1024$. This, however, doesn't imply that the overall performance of $\hat{K}_{NDA-LOSA}$ is better than that of $\hat{K}_{DA-LOSA}$ and that the overall performance of $\hat{K}_{NDA-LOSA}$ improves as the sample size decreases. Fig. 4.40 shows the performance of $\hat{K}_{NDA-LOSA}$ for different sample sizes. It is clear that the bias of $\hat{K}_{NDA-LOSA}$ will become negative and $\hat{K}_{NDA-LOSA}$ will underperform $\hat{K}_{DA-LOSA}$ again when K increases further. Also, when K increases further, $\hat{K}_{NDA-LOSA}$ will have a larger bias and a smaller standard deviation for small sample sizes than for large sample sizes. Since a larger bias with a smaller standard deviation means that the estimate is more concentrated on some value that is farther away from the true value, the overall performance of $\hat{K}_{NDA-LOSA}$ still deteriorates as L decreases. This phenomenon is probably caused by the time-variation of the fading mean, as $\hat{K}_{NDA-LOSB}$ in the constant LOS component model don't demonstrate such behavior.

Figs. 4.41 and 4.42 show the performances of $\hat{\bar{\rho}}_{DA-LOSB}$ and $\hat{\bar{\rho}}_{NDA-LOSB}$ for different values of

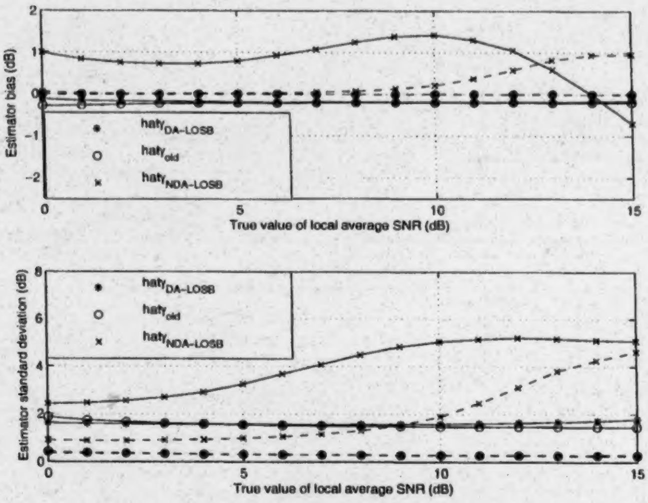

Figure 4.41. Performances of $\hat{\rho}_{DA-LOSB}$, $\hat{\rho}_{old}$ and $\hat{\rho}_{NDA-LOSB}$ with $L = 1024$ (dashed line) and $L = 128$ (solid line) at $K = 5$, assuming a constant LOS fading component.

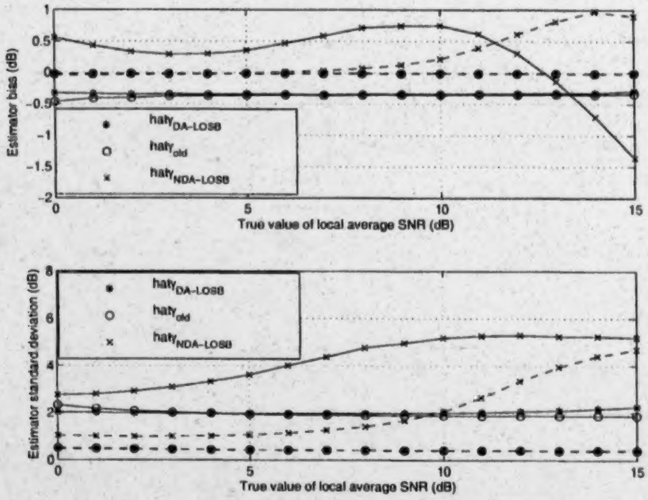

Figure 4.42. Performances of $\hat{\rho}_{DA-LOSB}$, $\hat{\rho}_{old}$ and $\hat{\rho}_{NDA-LOSB}$ with $L = 1024$ (dashed line) and $L = 128$ (solid line) at $K = 2.5$, assuming a constant LOS fading component.

K and L. One sees that $\hat{\rho}_{DA-LOSB}$ has a bias of about -0.1 dB and a standard deviation of about 1.8 dB when $L = 128$ and $K = 5$, a bias of about zero and a standard deviation of about 0.3 dB when $L = 1024$ and $K = 5$, a bias of about -0.3 dB and a standard deviation of about 2.0 dB when $L = 128$ and $K = 2.5$, and a bias of about zero and a standard deviation of about 0.3 dB when $L = 128$ and $K = 2.5$. The new estimator, $\hat{\rho}_{DA-LOSB}$, shows good performance. The estimator performance improves as the sample size increases, as expected. Comparing $\hat{\rho}_{DA-LOSB}$ with $\hat{\rho}_{NDA-LOSB}$, one observes that $\hat{\rho}_{DA-LOSB}$ is much better than $\hat{\rho}_{NDA-LOSB}$. Comparing $\hat{\rho}_{DA-LOSB}$ with $\hat{\rho}_{old}$, one sees that their performances are comparable when $L = 1024$. When $L = 128$, $\hat{\rho}_{DA-LOSB}$ slightly outperforms $\hat{\rho}_{old}$, as $\hat{\rho}_{DA-LOSB}$ has a smaller bias than $\hat{\rho}_{old}$. The new estimator, $\hat{\rho}_{DA-LOSB}$, benefits from taking the sample correlation into account, but it suffers from using estimates of two parameters P^2 and $2\alpha^2$. On the other hand, the previous estimator, $\hat{\rho}_{old}$, benefits from using an estimate of only one parameter $2\alpha^2$, but it suffers from not taking the sample correlation into account. When L is large, there are enough independent samples among the L correlated samples for $\hat{\rho}_{old}$ to obtain an accurate estimate of $2\alpha^2$, and the sample correlation doesn't dominate. In this case, they show comparable performances. When L is small such that few samples among the L correlated samples are independent, $\hat{\rho}_{old}$ cannot obtain an accurate estimate of $2\alpha^2$. In this case, the sample correlation dominates, and our correlated-sample-based estimator, $\hat{\rho}_{DA-LOSB}$, shows some advantage.

Figs. 4.43 and 4.44 show the performances of $\hat{\rho}_{DA-LOSA}$ and $\hat{\rho}_{NDA-LOSA}$ for different values of K and L. Similar observations can be made. Again, the DA estimator, $\hat{\rho}_{DA-LOSA}$, has good performance, and it outperforms the NDA estimator, $\hat{\rho}_{NDA-LOSA}$. Comparing $\hat{\rho}_{DA-LOSA}$ with $\hat{\rho}_{old}$, one sees that $\hat{\rho}_{DA-LOSA}$ is slightly better than $\hat{\rho}_{old}$ when $L = 128$. When $L = 1024$ and $\bar{\rho}$ is large, $\hat{\rho}_{old}$ is slightly better than $\hat{\rho}_{DA-LOSA}$, as the sample correlation is not dominant and our estimator $\hat{\rho}_{DA-LOSA}$ introduces extra estimation errors by using estimates of two parameters in this case; the performance difference is small.

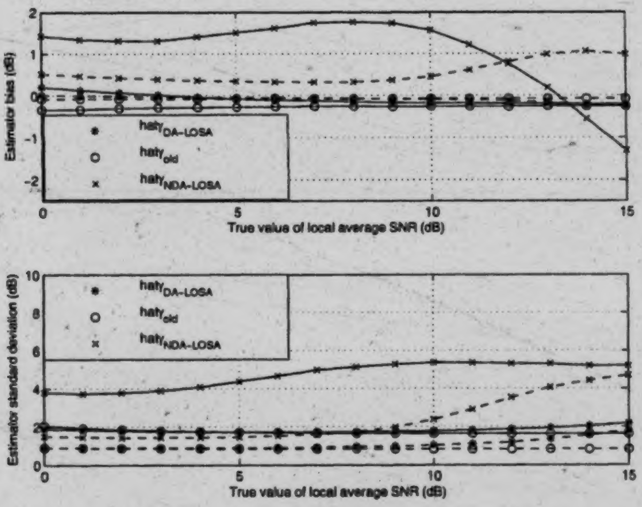

Figure 4.43. Performances of $\hat{\hat{\rho}}_{DA-LOSA}$, $\hat{\hat{\rho}}_{old}$ and $\hat{\hat{\rho}}_{NDA-LOSA}$ with $L = 1024$ (dashed line) and $L = 128$ (solid line) at $K = 5$, assuming a time-varying LOS fading component.

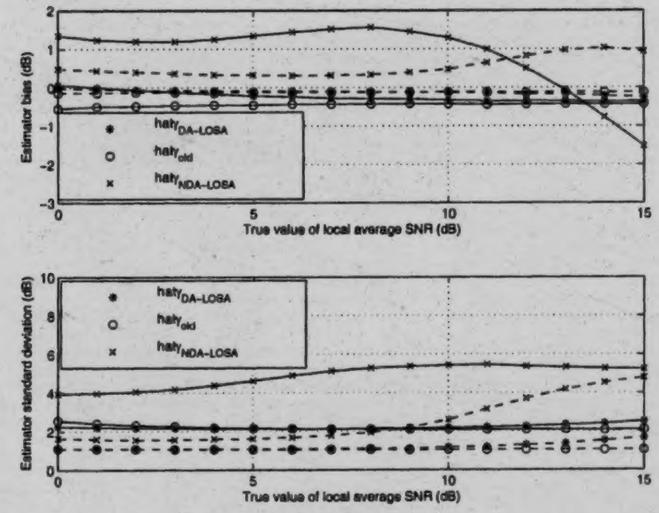

Figure 4.44. Performances of $\hat{\hat{\rho}}_{DA-LOSA}$, $\hat{\hat{\rho}}_{old}$ in and $\hat{\hat{\rho}}_{NDA-LOSA}$ with $L = 1024$ (dashed line) and $L = 128$ (solid line) at $K = 2.5$, assuming a time-varying LOS fading component.

4.5 Moment-based Estimation of SINR

In the previous sections, estimation of SNR and ASNR has been performed. The estimation is based on the assumption of no interferences. In some wireless communication systems, interferences may also occur, in addition to the noise. In this case, the SINR, defined as the ratio of the signal power to the interference-plus-noise power, is frequently used as an important link quality measure. Estimation of SINR has been studied previously by several researchers [100]- [104]. These estimators assume either known data symbols or constant modulus signaling. However, for a high data rate transmission system using QAM, known data symbols may not be available in the signaling format or may be undesirable since they reduce throughput. New estimators that blindly estimate the SINR of QAM signals are desired. In [105] and [106], non-data-aided (NDA) SNR estimators were derived for QAM signals in an additive white Gaussian noise channel and a slowly fading channel, respectively. These estimators are designed to operate in the absence of interferences. In this work, we examine the more general case when interference and noise are both present. We derive NDA SINR estimators for QAM signals in a slowly fading channel under two channel conditions.

4.5.1 Channel Condition 1

First, we consider channels where the interference can be modeled as a Gaussian random variable with non-zero mean. The received signal is

$$y_l = As_l + I_l + n_l \tag{4.87}$$

where $l = 1, 2, \cdots, L$ index different samples, A is the unknown complex channel gain assumed constant over the estimation time in a slowly fading channel, s_l is the unknown transmitted signal in the l-th sample, I_l is additive Gaussian interference with mean e and variance $2\alpha^2$, and n_l is additive Gaussian noise with mean zero and variance $2\sigma^2$. Note that the channel gain of the desired signal is subjected to fading, but the fading is assumed constant over the estimation time [100]- [106]. This can be achieved by properly selecting the estimation time. For example, when the channel is fading slowly enough such that the channel gain remains approximately the same over one packet period (but it may change from one packet to another), the estimation can be performed over the time duration of one packet and the fading channel gain is constant for all samples obtained during this time period. This assumption applies to each of the interfering signals as well. However, the overall interference is modeled as a Gaussian random variable [14], [107], which differs from sample to sample, similar to the noise. Assume that the average energy of the transmitted signal, $E\{|s_l|^2\}$, is normalized to 1. Also, assume that the interference and the noise are circulary symmetric with independent real and imaginary parts, and that the noise is independent of the interference. Denote $c_l = |I_l|$ as the amplitude of the interference. Then, c_l is a Ricean random variable with probability density function (PDF) $f_{c_l}(x) = \frac{2(K+1)x}{\Omega} e^{-K-\frac{(K+1)x^2}{\Omega}} I_0\left(2\sqrt{\frac{K(K+1)}{\Omega}}x\right)$, where $K = \frac{|e|^2}{2\alpha^2}$ is the Ricean K factor and $\Omega = |e|^2 + 2\alpha^2$ is the total mean power of the interference. The interference model in (4.87) is flexible. For example, when $K = 0$ or $e = 0$, I_l becomes a zero mean Gaussian random variable. This represents channels where there are many weak interferers plus noise but no dominant interferers. When $K \to \infty$ or $\alpha^2 = 0$, I_l becomes a constant. This represents channels where there is one dominant constant interferer plus noise throughout the estimation. Finally, when the value of K is between 0 and ∞, this represents channels where there are many weak interferers, a strong but not dominant constant interferer, and noise. The SINR to be estimated in this case can be defined as $\rho_1 = \frac{|A|^2}{\Omega+2\sigma^2}$. Next, we derive a moment-based estimator for ρ_1.

From (4.87), one has

$$E\{y_l\} = e \tag{4.88}$$

$$E\{|y_l|^2\} = |A|^2 + \Omega + 2\sigma^2 \tag{4.89}$$

and

$$E\{|y_l|^4\} = P|A|^4 + 2(\Omega + 2\sigma^2)^2 + 4|A|^2(\Omega + 2\sigma^2) - |e|^4 \tag{4.90}$$

where $P = E\{|s_l|^4\}$ and $E\{s_l\} = 0$ for most practical QAM signalings. Using (4.88), (4.89) and (4.90), one has an equation for ρ_1 as $h_1 = \frac{P\rho_1^2+4\rho_1+2}{\rho_1^2+2\rho_1+1}$, where $h_1 = \frac{E\{|y_l|^4\}+|E\{y_l\}|^4}{E^2\{|y_l|^2\}}$. Solving this equation, one moment-based estimator for ρ_1 can be derived as

$$\hat{\rho}_1 = \frac{-b_1 + \sqrt{b_1^2 - 4a_1c_1}}{2a_1} \tag{4.91}$$

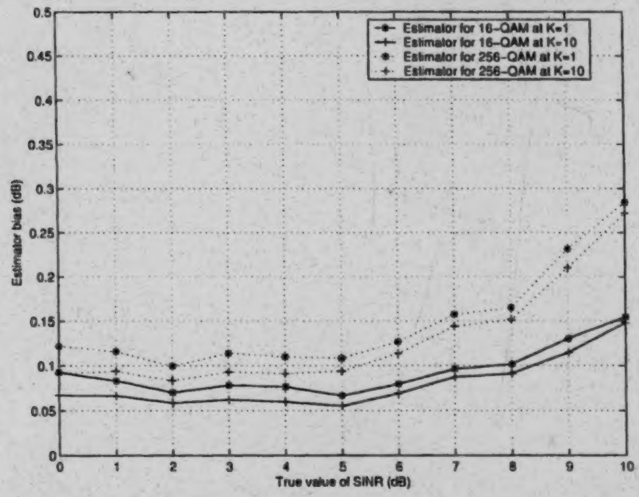

Figure 4.45. The biases of $\hat{\rho}_1$ for 16-QAM and 256-QAM signals for different values of K.

where $a_1 = \hat{h}_1 - P$, $b_1 = 2\hat{h}_1 - 4$, $c_1 = \hat{h}_1 - 2$, $\hat{h}_1 = \frac{\hat{\mu}_4 + |\hat{\mu}_1|^4}{\hat{\mu}_2^2}$, $\hat{\mu}_1 = \frac{1}{L}\sum_{l=1}^{L} y_l$, $\hat{\mu}_2 = \frac{1}{L}\sum_{l=1}^{L} |y_l|^2$, $\hat{\mu}_4 = \frac{1}{L}\sum_{l=1}^{L} |y_l|^4$, and $y_1, y_2, \cdots, y_L$ are L independent and identically distributed samples.

Figs. 4.45 and 4.46 show the bias and root mean squared error (RMSE) of $\hat{\rho}_1$, respectively. Two signaling schemes, 16-QAM and 256-QAM, are examined. Similar to [106], a sample size of $L = 1000$ is used. Channel conditions of $K = 1$ and $K = 10$ are considered. One sees that the bias of $\hat{\rho}_1$ is between $0.05dB$ and $0.15dB$ for 16-QAM and between $0.09dB$ and $0.28dB$ for 256-QAM, and the RMSE of $\hat{\rho}_1$ is between 0.25 and 2.20 for 16-QAM and between 0.30 and 3.20 for 256-QAM, over all the values of SINR considered. Therefore, the estimator $\hat{\rho}_1$ performs well in this SINR region. The bias and the RMSE of $\hat{\rho}_1$ increase as the true value of SINR increases. A larger constellation size corresponds to a larger bias as well as a larger RMSE. Also, comparing the performance of $\hat{\rho}_1$ at $K = 1$ with that at $K = 10$, one sees that the estimate has a smaller bias and smaller RMSE in a channel with larger K. These observations suggest that $\hat{\rho}_1$ works better for QAM with smaller constellation size in a channel with larger K.

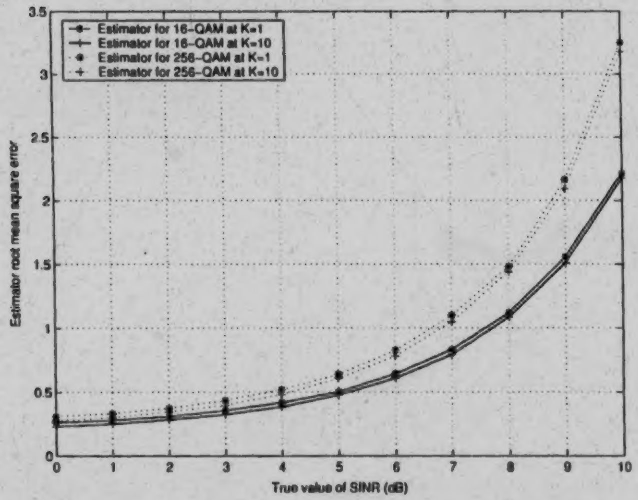

Figure 4.46. The root mean squared errors of $\hat{\rho}_1$ for 16-QAM and 256-QAM signals for different values of K.

4.5.2 Channel Condition 2

Next, we consider channels where there is one dominant synchronous interferer. The received signal in this case can be written as

$$y_l = As_l + Bt_l + n_l \tag{4.92}$$

where B is the channel gain of the dominant synchronous interfering user affecting the desired user, t_l is the transmitted signal component of the dominant synchronous interfering user in the l-th sample, and l, A, s_l and n_l are defined as before. Note that this case cannot be modeled by (4.87). Note further that the desired signal and the interfering signal have the same statistical structures in this case. For simplicity, we assume that t_l has the same signaling scheme as s_l. Extension to the case when s_l and t_l have different signaling schemes is straightforward. The SINR to be estimated in this case can be defined as $\rho_2 = \frac{|A|^2}{|B|^2+2\sigma^2}$. From (4.92), one has

$$E\{|y_l|^2\} = |A|^2 + |B|^2 + 2\sigma^2 \tag{4.93}$$

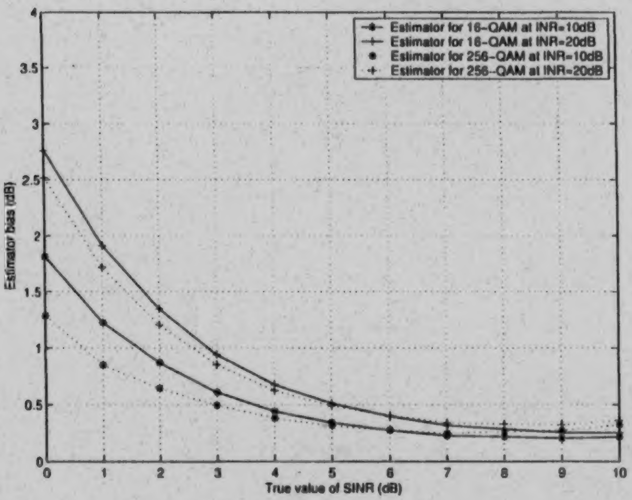

Figure 4.47. The biases of $\hat{\rho}_2$ for 16-QAM and 256-QAM signals for different values of INR.

and

$$E\{|y_l|^4\} = P|A|^4 + 2(|B|^2 + 2\sigma^2)^2 + 4|A|^2(|B|^2 + 2\sigma^2) + (P-2)|B|^4. \tag{4.94}$$

Using (4.93) and (4.94), one has an equation for ρ_2 as $h_2 = \frac{P\rho_2^2+4\rho_2+d}{\rho_2^2+2\rho_2+1}$, where $d = 2 + \frac{P-2}{(1+\frac{1}{INR})^2}$, $h_2 = \frac{E\{|y_l|^4\}}{E^2\{|y_l|^2\}}$ and $INR = \frac{|B|^2}{2\sigma^2}$ is the interference-to-noise ratio (INR). When $INR \gg 1$, one has $d \approx P$, and P is around 1.3 for 16-QAM signaling. When $INR \ll 1$, one has $d \approx 2.0$. After testing several values of d between 1.3 and 2.0 by simulation, we found that $d = 1.6$ gives a SINR estimator with best overall performance in the cases considered. Thus,

$$h_2 \approx \frac{P\rho_2^2 + 4\rho_2 + 1.6}{\rho_2^2 + 2\rho_2 + 1}. \tag{4.95}$$

By solving (4.95), one has a moment-based estimator for ρ_2 as

$$\hat{\rho}_2 = \frac{-b_2 + \sqrt{b_2^2 - 4a_2c_2}}{2a_2} \tag{4.96}$$

where $a_2 = \hat{h}_2 - P$, $b_2 = 2\hat{h}_2 - 4$, $c_2 = \hat{h}_2 - 1.6$, and $\hat{h}_2 = \frac{\hat{\mu}_4}{\hat{\mu}_2^2}$.

Figs. 4.47 and 4.48 show the bias and RMSE of $\hat{\rho}_2$, respectively. Again, two signaling schemes, 16-QAM and 256-QAM, are examined. We consider the cases when $INR = 10dB$ and $INR = 20dB$.

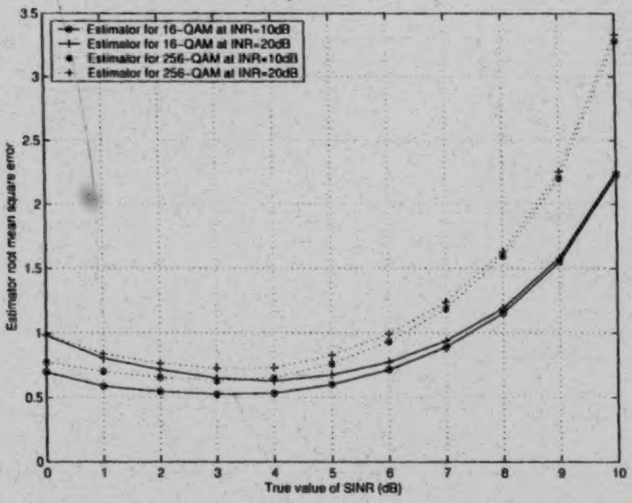

Figure 4.48. The root mean squared errors of $\hat{\rho}_2$ for 16-QAM and 256-QAM signals for different values of INR.

The estimator $\hat{\rho}_2$ has a bias between 0.2 dB and 2.7 dB for 16-QAM and between 0.3 dB and 2.5 dB for 256-QAM, and a RMSE between 0.5 and 2.3 for 16-QAM and between 0.6 and 3.3 for 256-QAM, in the SINR region considered. The bias of $\hat{\rho}_2$ decreases as the true value of SINR increases and becomes approximately constant at large values of SINR. The RMSE of $\hat{\rho}_2$ decreases as the SINR increases over the range of 0 dB to 3 dB, and it increases as the SINR increases over the range of 3 dB to 10 dB. One also sees that $\hat{\rho}_2$ for 16-QAM has a smaller RMSE than $\hat{\rho}_2$ for 256-QAM, over all the values of SINR considered. The bias of $\hat{\rho}_2$ for 16-QAM is comparable to that for 256-QAM. Finally, observe that the estimate has a smaller bias and smaller RMSE in a channel with smaller INR. These observations also suggest that $\hat{\rho}_2$ works better for QAM with smaller constellation size in a channel with smaller INR. The above results consider QAM signals only. However, they can be extended to constant modulus signals as well.

Chapter 5

Conclusions and Future Work

In this chapter, we first highlight important findings of this thesis, and then suggest several topics for future research.

5.1 Conclusions

1. Maximum likelihood decision-based estimators for the channel state parameter in a static AWGN channel and a slowly fading channel have been developed. The approximations to the probability density functions of the maximum likelihood estimates have been derived under the assumption of no decision errors. Based on these approximate functions, the estimator performances have been examined and possible performance improvement has been stated.
2. The performances of maximum estimated branch signal amplitude SDC and maximum estimated branch SNR SDC using practical channel estimators have been analyzed, under the assumption of independent and non-identically distributed diversity branches and unequal noise powers. The effects of channel estimation errors and imbalanced noise powers have been quantified in terms of closed-from expressions for the error rates.
3. When the noise powers are balanced, maximum estimated branch signal amplitude SDC performs better than maximum estimated branch SNR SDC. When the noise powers are imbalanced, maximum estimated branch SNR SDC performs better than maximum estimated

branch signal amplitude SDC.

4. Novel diversity receivers that operate in the presence of Gaussian channel estimation errors have been proposed by using statistics of channel estimation errors. The structures of the novel receivers depend on the nature and the amount of knowledge of error statistics available.
5. The conventional MRC receiver is suboptimal when channel estimation errors occur. The new diversity receivers outperform the conventional MRC receiver in most cases considered.
6. Optimum PSAM signal detectors for Rayleigh and Ricean fading channels have been derived. The conventional PSAM signal detector is proved to be optimum in the sense of minimum probability of error for BPSK signaling in Rayleigh fading channels. It is suboptimal for 16-QAM signaling in Rayleigh fading channels and for BPSK signaling and 16-QAM signaling in Ricean fading channels.
7. The performance gain of the optimum PSAM signal detector for BPSK signaling comes from the use of the specular component in Ricean fading channels, while the performance gain of the optimum PSAM signal detector for 16-QAM signaling comes from the use of specular component in Ricean fading channels as well as the joint processing of data and pilot symbols.
8. The CRLBs for NDA ML channel gain and delay estimation in an ultra-wide bandwidth system have been derived analytically. Novel NDA ML estimators for the channel gain and delay in an ultra-wide bandwidth system have also been designed by using a more accurate approximation to the log-likelihood function.
9. Noisy sample based ML and moment-based estimators for the fading distribution parameters have been derived, under the assumption of known or unknown noise power. These estimators have superior performances when operating in a practical noisy fading channel. They are very useful for channel modeling and receiver design in realistic communication systems.

10. Novel estimators for the Ricean K parameter have been derived by using fading phase samples. The new estimators have much simpler structures than previous envelope-based estimators. In most cases, they also have better performances than the previous estimators.

11. Maximum likelihood estimators for two measures of SNR in a static AWGN channel and a slowly fading channel have been proposed, considering both sampled signal system and continuous signal system. The approximations to the probability density functions of the SNR estimates have been derived under the assumption of no decision errors.

12. The ML estimator for SNR is a decision-based structure that necessarily incorporates a digital data detector. Therefore, the optimal performance of ML estimation of SNR can be achieved at minimal additional cost in a digital signal receiver.

13. A novel approximate ML estimator for SNR in a static AWGN channel has been derived by using all symbols in a frame. This estimator can be applied to any system where signals are transmitted in frames and the frame has both known and unknown fields.

14. Maximum likelihood estimator for the average SNR in a Ricean fading channel has been derived. The probability density function of the estimate has been obtained, and the mean, the variance and the mean squared error have been calculated to show its asymptotical optimality.

15. Joint estimation of the K parameter and the ASNR in a Ricean fading channel has been performed. New estimators for the Ricean K parameter and the ASNR have been designed by using noisy and correlated channel samples.

16. Non-data-aided moment-based estimators for the SINR in a system using QAM signals have been derived, considering two different practical channel conditions.

5.2 Future Work

1. The performances of the ML SNR estimators in a static AWGN channel and a slowly fading channel are analyzed under the assumption of no decision errors. This analysis is not valid for

decision-based estimation with low operating SNR. In some cases, it is of interest to analyze the performances of the decision-based ML SNR estimators with decision errors.

2. The proposed ML SNR estimators can be used in the acquisition mode of the system. It is also of interest to extend these results to the tracking mode of the system, where one uses current received signals to update the SNR estimate continuously. In this way, the estimator adapts to the channel changes. Another useful extension is to design an iterative algorithm where the decision-based SNR estimate is fed back into the detector to improve the data decision and the improved data decision is used for SNR estimation again in the next iteration. In this way, both the SNR estimation and the data decision may be improved.

3. The values of statistics of the channel estimation errors are assumed known in the receiver design. In practice, they have to be obtained by estimation as well. There will be a mismatch between the true value and the estimate of the error statistics. It is of interest to evaluate the effect of this mismatch on the receiver performance.

4. The proposed optimum PSAM signal detector doesn't take diversity combining into account. It is well known that diversity combining is a very effective method in combating fading. Therefore, it is of great interest to investigate the problem of optimum PSAM signal detection in a diversity system.

5. The fading distribution parameters have been estimated by using practical noisy channel samples. However, these estimators don't consider interferences in the system. Interference-limited systems, such as those code division multiple access systems, have been widely used currently. Design of fading distribution parameter estimators in an interference-limited system is still a challenging problem.

6. The results in the thesis consider flat fading channels and single-input systems only. It is also of interest to extend them to a frequency-selective fading channel and a multiple-input-multiple-output system.

References

[1] J.G.Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.

[2] M. Abramowitz and I.A. Stegun, Eds. *Handbook of Mathematical Functions*. New York: Dover, 1972.

[3] M.K. Simon, S.M. Hinedi and W.C. Lindsey, *Digital Communications Techniques: Signal Design and Detection*. Upper Saddle River, NJ: Prentice Hall, 1995.

[4] M. Schwartz, W.R. Bennet, and S. Stein, *Communication Systems and Techniques*. New York: IEEE Press, reprint 1995, 1966.

[5] J.C. Hancock and W.C. Lindsey, "Optimum performance of self-adaptive systems operating through a Rayleigh fading medium," *IEEE Trans. Commun. Sys.*, vol. 11, pp. 443-453, Dec. 1963.

[6] U. Charash, "Reception through Nakagami fading multipath channels with random delays," *IEEE Trans. Commun.*, vol. 27, pp. 657-670, Apr. 1979.

[7] M.K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels*, 2nd Ed., New York: John Wiley & Sons, 2004.

[8] R. Steele and L. Hanzo, Eds. *Mobile Radio Communications: Second and Third Generation Cellular and WATM Systems*, 2nd Ed., New York: John Wiley & Sons, 1999.

[9] C. Tepedelenlioglu, A. Abdi, G. Giannakis, " The Ricean K factor: estimation and performance analysis," *IEEE Trans. Wireless Commun.*, vol. 2, pp. 799-810, July 2003.

[10] Y. Ko and M.-S. Alouini, "Estimation of Nakagami fading channel parameters with application to optimized transmitter diversity systems," *Pro. ICC 2001*, Helsinki, Finland, June 2001.

[11] J. Zander, "Performance of optimum transmitter power control in cellular radio systems," *IEEE Trans. Veh. Technol.*, vol. 41, pp. 57-62, Feb. 1992.

[12] E.A. Newcombe and S. Pasupathy, "Error rate monitoring for digital communications," *Proc. IEEE*, vol. 70, pp. 805-828, Aug. 1982.

[13] K. Balachandran, S.R. Kadaba and S. Nanda, " Channel quality estimation and rate adaptation for cellular mobile radio," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1244-1256, July 1999.

[14] H. Vincent Poor, *An Introduction to Signal Detection and Estimation*, 2nd Ed. New York: Springer, 1994.

[15] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ: Prentice Hall, 1993.

[16] K.S. Shanmugan and A.M. Breipohl, *Random Signals: Detection, Estimation and Data Analysis*. New York: Wiley, 1988.

[17] H.L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*. New York: Wiley, 2001.

[18] T.S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Upper Saddle River: Prentice-Hall, 2002.

[19] R.B. Kerr, "On signal and noise level estimation in a coherent PCM channel," *IEEE Trans. Aerop. Electron. Syst.*, vol. AES-2, pp. 450-454, July 1966.

[20] R.M. Gagliardi and C.M. Thomas, "PCM data reliability monitoring through estimation of signal-to-noise ratio," *IEEE Trans. Commun.*, vol. COM-16, pp. 479-486, June 1968.

[21] C.M. Thomas, "Maximum Likelihood Estimation of Signal-to-Noise Ratio," Ph.D. dissertation, Univ. of Southern California, Los Angeles, 1967.

[22] D.R. Pauluzzi and N.C. Beaulieu, "A comparison of SNR estimation techniques for the AWGN channel," *IEEE Trans. Commun.*, vol. 48, pp. 1681-1691, Oct. 2000.

[23] N.C. Beaulieu, A.S. Toms and D.R. Pauluzzi, "Comparison of four SNR estimators for QPSK modulations," *IEEE Commun. Lett.*, vol. 4, pp. 43-45, Feb. 2000.

[24] S. Yoshida, GIL. Tan, H. Zhou, and T. Takeuchi, "Simple method of multipath delay difference detection for $\pi/4$-shift QPSK," *Electron, Lett.*, vol. 27, pp. 1027-1028, June 1991.

[25] B. Li, R. DiFazio and A. Zeira, "A low bias algorithm to estimate negative SNRs in an AWGN channel," *IEEE Commun. Lett.*, vol. 6, pp. 469-471, Nov. 2002.

[26] N.S. Alagha, "Cramer-Rao bounds of SNR estimates for BPSK and QPSK modulated signals," *IEEE Commun. Lett.*, vol 5, pp. 10-12, Jan. 2001.

[27] A. Wiesel, J. Goldberg and H. Messer, "Non-data-aided signal-to-noise estimation," *Proc. ICC 2002*, vol. 1, pp. 197-201, 2002.

[28] M. Schwartz and L. Shaw, *Signal Processing: Discrete Spectral Analysis, Detection and Estimation.* New York: McGraw-Hill, 1975.

[29] N.C. Beaulieu and C. Leung, "On the performance of three suboptimum detection schemes for binary signaling," *IEEE Trans. Commun.*, vol. COM-33, pp. 241-245, Mar. 1985.

[30] W. Rudin, *Real and Complex Analysis.* New York: McGraw-Hill, 1966.

[31] S.O. Rice, "Mathematical analysis of random noise," Bell. Sys. Tech. J., vol. 23, pp. 282-332, July 1944; vol. 24, pp. 46-156, Jan. 1945; also in *Selected Papers on Noise and Stochastic Processes*, N. Wax, Ed. New York: Dover, 1954.

[32] M.I. Schwartz, "Distribution of the time-average power of a Gaussian process," *IEEE Trans. Inform. Theory*, vol. 16, pp. 17-26, 1970.

[33] N.C. Beaulieu, "Closed-form approximation to PDF of time-average power of bandlimited Gaussian process," *Electron. Lett.*, vol. 32, pp. 1948-1950, Oct. 1996.

[34] A. Papoulis and S.U. Pillai, *Probability, Random Variables and Stochastic Processes*, 4th ed. New York: McGraw-Hill, 2002.

[35] I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series, and Products*, 6th ed. San Diego: Academic Press, 2000.

[36] M. Umehira, A. Ohta, S. Kagami, H. Hojo, and T. Kobayashi, "A 5 GHz-band advanced wireless access system for mobile multimedia applciations," *Proc. IEEE Vehicular Technol. Conference (VTC 2000)*, vol. 3, pp. 2300-2304, May 2000.

[37] E. Grayver and B. Daneshrad, "A low-power all-digital FSK receiver for space applications," *IEEE Trans. Commun.*, vol. 49, pp. 911-921, May 2001.

[38] E. Grayver and B. Daneshrad, "VLSI implementation of a 100-μ W multirate FSK receiver," *IEEE J. Solid-State Circuits*, vol. 36, pp. 1821-1828, Nov. 2001.

[39] A. Annamalai Jr., "The effect of Gaussian error in selection diversity combiners", *Wirel. Commun. Mob. Comput.*, vol. 1, pp. 419-435, Oct./Dec. 2001.

[40] Y. Ma, R. Schober and S. Pasupathy, "Performance of M-PSK with GSC and EGC with Gaussian weithing errors," *IEEE Trans. Vehicul. Technol.*, vol. 54, pp. 149-162, Jan. 2005.

[41] P. Shamain and L.B. Milstein, "Effect of mutual coupling and correlated fading on receive diversity systems using compact antenna arrays and noisy channel estimates," *Proc. IEEE Globecom 2003*, vol. 3, pp. 1669-1673, Dec. 2003.

[42] P. Lombardo and D. Pastina, "Quiescent pattern control in adaptive antenna processing at subarray level," *Proc. IEEE International Symposium on Phased Array Systems and Technology*, pp. 176-181, Oct. 2003.

[43] M. Chiani, M.Z. Win and A. Zanella, "Error probability for optimum combining of M-ary PSK signals in the presence of interference and noise," *IEEE Trans. Commun.*, vol. 51, pp. 1949-1957, Nov. 2003.

[44] A. Aghamohammadi and H. Meyr, "On the error probability of linearly modulated signals on Rayleigh frequency-flat fading channels," *IEEE Trans. Commun.*, vol. 38, pp. 1966-1970, Nov. 1990.

[45] M. Gans, "The effect of Gaussian error in maximal ratio combiners", *IEEE Trans. Commun.*, vol. COM-19, pp. 492-500, Aug. 1971.

[46] W.B. Davenport and W.L. Root, *An Introduction to the Theory of Random Signals and Noise*. New York: McGraw-Hill, 1958.

[47] L. Cao and N.C. Beaulieu, "Exact error-rate analysis of diversity 16-QAM with channel estimation error," *IEEE Trans. Commun.*, vol. 52, pp. 1019-1029, June 2004.

[48] L. Cao and N.C. Beaulieu, "Closed-form results for the BER of maximal ratio combining with channel estimation errors in Ricean fading," to appear in *IEEE Trans. Wireless Commun.*.

[49] N.L. Johnson and N. Balakrishnan, *Continuous Univariate Distributions: vol. 2*, 2nd Ed. New York: John Wiley & Sons, 1995.

[50] A. Jeffrey, *Handbook of Mathematical Formulas and Integrals*, 2nd ed. San Diego: Academic Press, 2000.

[51] F. Ling, "Optimal reception, performance bound, and cutoff rate analysis of references-assisted coherent CDMA communications with applications," *IEEE Trans. Commun.*, vol. 47, pp. 1583-1592, Oct. 1999.

[52] W. Xiao, "Optimal detection of M-QAM signal with channel estimation error," in *Proc. IEEE International Conf. on Commun. (ICC'03)*, pp. 3251-3255, May 2003.

[53] M.M. Wang, W. Xiao, and T. Brown, "Soft decision metric generation for QAM with channel estimation error," *IEEE Trans. Commun.*, vol. 50, pp. 1058-1061, July 2002.

[54] G. Stuber, *Principles of Mobile Communications*, 2nd Ed., Boston: Kluwer Academic, 2001.

[55] J.K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Vehicul. Technol.*, vol. 42, pp. 686-693, Nov. 1991.

[56] S. Sampei and T. Sunaga, "Rayleigh fading compensation for QAM in land mobile radio communications," *IEEE Trans. Vehicul. Technol.*, vol. 42, pp. 137-147, May 1993.

[57] H. Li, Y. Iwanami, and T. Ikeda, "Symbol error rate analysis for MPSK under Rician fading channels with fading compensation based on time correlation," *IEEE Trans. Vehicul. Technol.*, Vol. 44, pp. 535-542, Aug. 1995.

[58] J.G. Proakis, "Probabilities of error for adaptive reception of M-phase signals," *IEEE Trans. Commun. Technol.*, vol. COM-16, pp. 71-81, Feb. 1968.

[59] Y. Chen and N.C. Beaulieu, "Estimators using noisy channel samples for fading distribution parameters,", *IEEE Trans. Commun.*, vol. 53, pp. 1274- 1277, Aug. 2005.

[60] M.Z. Win and R.A. Scholtz, "Ultra-wide bandwidth time-hopping spread-spectrum impulse radio for wireless multiple-access communications," *IEEE Trans. Commun.*, vol. 48, pp. 679-691, Apr. 2000.

[61] M.Z. Win and R.A. Scholtz, "Impulse radio: how it works," *IEEE Commun. Lett.*, vol. 2, pp. 36-38, Feb. 1998.

[62] S. Roy, J.R. Foerster, V.S. Somayazulu, and D.G. Leeper, "Ultra-wideband radio design: the promise of high-speed, short-range wireless connectivity," *Proc. IEEE*, vol. 92, pp. 295-311, Feb. 2004.

[63] M.Z. Win and R.A. Scholtz, "On the energy capture of ultrawide bandwidth signals in dense multipath environments," *IEEE Commun. Lett.*, vol. 2, pp. 245-247, Sept. 1998.

[64] V. Lottici, A. D'Andrea, and U. Mengali, "Channel estimation for ultra-wideband communications," *IEEE J. Select. Areas Commun.*, vol. 20, pp. 1638-1645, Dec. 2002.

[65] L. Huang and C.C. Ko, "Performance of maximum-likelihood channel estimator for UWB communcations," *IEEE Commun. Lett.*, vol. 8, pp. 356-358, June 2004.

[66] B. Hu and N.C. Beaulieu, "Pulse shaping in UWB communication systems," *Proc. IEEE Vehicular Technology Conference - Fall*, Los Angeles, Sept. 26-29, 2004.

[67] P.T. Nielsen, "Mean power of a Rice fading signal determined from measurements," *Proc. IEEE*, vol. 66, pp. 1094-1095, Sept. 1978.

[68] A. Abdi, C. Tepedelenlioglu, M. Kaveh and G. Giannakis, " On the estimation of the K parameter for the Rice fading distribution", *IEEE Comm. Lett.*, vol. 5, pp. 92-94, Mar. 2001.

[69] L.J. Greenstein, D.G. Michelson, and V. Erceg, "Moment-based estimation of the Ricean K-factor," *IEEE Commun. Lett.*, vol. 3, pp. 175-176, June 1999.

[70] K.K. Talukdar, W.D. Lawing, "Estimation of the parameters of the Rice distribution," *J. Acoust. Soc. Am.*, vol. 89, pp. 1193-1197, Mar. 1991.

[71] G. Azemi, B. Senadji, and B. Boashash, "Ricean K-factor estimation in mobile communication systems," *IEEE Commun. Lett.*, vol. 8, pp. 617-619, Oct. 2004.

[72] G. Azemi, B. Sedadji, and B. Boashash, "Estimating the Ricean K-factor for mobile communication applications," *Proc. ISSPA'03*, vol. 2, pp. 311-314, Paris, France, July 2003.

[73] M. Nakagami, "The m-distribution - a general formula of intensity distribution of rapid fading," in *Statistical Method in Radio Wave Propagation*, W.C. Hoffman Ed., Oxford, U.K.: Pergamon, pp. 3-36, 1960.

[74] J. Cheng and N.C. Beaulieu, " Generalized moment estimators for the Nakagami fading parameter," *IEEE Comm. Lett.*, vol. 6, pp. 144-146, Apr. 2002.

[75] J.Cheng and N.C.Beaulieu, "Maximum-likelihood based estimation of the Nakagami m parameter," *IEEE Commun. Lett.*, vol. 5, pp. 101-103, Mar. 2001.

[76] J.Cheng and N.C.Beaulieu, "Moment-based estimation of the Nakagami-m fading parameter," *IEEE PACRIM Conf. on Comm. Comp. and Signal Proc.*, vol. 2, pp. 361-364, 2001.

[77] Q.T. Zhang, " A note on the estimation of Nakagami-m fading parameter," *IEEE Comm. Lett.*, vol. 6, pp. 237-238, June 2002.

[78] A. Abdi and M. Kaveh, "Performance comparison of three different estimators for the Nakagami m paramter using Monte Carlo simulation," *IEEE Commun. Lett.*, vol. 4, pp. 119-121, April 2000.

[79] C.Tepedelenlioglu, "Analytical performance analysis of moment-based estimators of the Nakagami parameter,", *IEEE Conf. on Vehicular Technology*, vol. 3, pp. 1471-1474, 2002.

[80] C. Tepedelenlioglu and P. Gao, "Estimators of the Nakagami-m parameter and performance analysis," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 519-527, Mar. 2005.

[81] M.M. Peritsky, "Statistical estimation of mean signal strength in a Rayleigh-fading environment," *IEEE Trans. Commun.*, vol. COM-21, pp. 1207-1213, Nov. 1973.

[82] N.C. Beaulieu, "Generation of correlated Rayleigh fading envelopes," *IEEE Commun. Lett.*, vol. 3, pp. 172-174, June 1999.

[83] J.D. Parsons, *The Mobile Radio Propagation Channel*, 2nd Ed., New York: John Wiley & Sons, 2000

[84] A.J. Viterbi, *CDMA: Principles of Spread Spectrum Communication*, Reading, MA: Addison-Wesley, 1998.

[85] M.K. Simon and M.-S. Alouini, "Average bit-error probability performance for optimum diversity combining of noncoherent FSK over Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 51, pp. 566-569, Apr. 2003.

[86] A.A. Abu-Dayya and N.C. Beaulieu, "Micro- and macrodiversity NCFSK (DPSK) on shadowed Nakagami-fading channels," *IEEE Trans. Commun.*, vol. COM-42, pp. 2693-2702, Sept. 1994.

[87] E.K. Hall and S.G. Wilson, "Design and analysis of turbo codes on Rayleigh fading channels," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 160-174, Feb. 1998.

[88] D.G. Brennan, "Linear diversity combining techniques," *Proc. IRE*, vol. 47, pp. 1075-1102, June 1959.

[89] D. Wong and D.C. Cox, "Estimating local mean signal power level in a Rayleigh fading environment," *IEEE Trans. Vehicular Technol.*, vol. 48, pp. 956-959, May 1999.

[90] Y.-C. Ko and T. Luo, "SNR estimation in decibel domain over communication channels," *Proc. IEEE Wireless Commun. and Networking Conf. (WCNC'04)*, pp. 1-6, Mar. 2004.

[91] A. Ramesh, A. Chockalingam, and L.B. Milstein, "Performance of noncoherent turbo detection on Rayleigh fading channels," *Proc. IEEE Global Telecommun. Conf. (Globecom'01)*, pp. 1193-1198, Nov. 2001.

[92] A. Ramesh, A. Chockalingam, and L.B. Milstein, "SNR estimation in Nakagami-*m* fading with diversity combining and its application to turbo decoding," *IEEE Trans. Commun.*, vol. 50, pp. 1719-1724, Nov. 2002.

[93] C.E. Gilchriest, "Signal-to-noise monitoring," *JPL Space Programs Summary*, vol. IV, nl. 3737-27, pp. 169-184, June 1966.

[94] "Physical Layer Standard for cdma2000 spread spectrum Systems - Revision D," 3GPP2 TSG-C C.S0002-D v1.0, Mar. 2004.

[95] 3GPP TS.25.211, *Physical channels and mapping of transport channels onto physical channels (FDD)*, Release 6, www.3gpp.org.

[96] K.D. Anim-Appiah, "On generalized covariance-based velocity estimation," *IEEE Trans. Vehicular Technol.*, vol. 48, pp. 1546-1557, Sept. 1999.

[97] J.M. Holtzman and A. Sampath, "Adaptive averaging methodology for handoffs in cellular systems," *IEEE Trans. Vehicular Technol.*, vol. 44, pp. 59-66, Feb. 1995.

[98] R. Narasimhan and D.C. Cox, "Speed estimation in wireless systems using wavelets," *IEEE Trans. Commun.*, vol. 47, pp. 1357-1364, Sept. 1999.

[99] L. Zhao and J.W. Mark, "Mobile speed estimation based on average fade slope duration," *IEEE Trans. Commun.*, vol. 52, pp. 2066-2069, Dec. 2004.

[100] A.L. Brandao, L.B. Lopez and D.C. McLernon, "Cochannel interference estimation for M-ary PSK modulated signals," *Kluwer Wireless Personal Commun.*, vol. 1, pp. 23-32, 1994.

[101] M.D. Austin and G.L. Stuber, "In-service signal quality estimation for TDMA cellular systems," *Kluwer Wireless Personal Commun.*, vol. 2, pp. 245-254, 1995.

[102] M. Turkboylari and G.L. Stuber, "An efficient algorithm for estimating the signal-to-interference ratio in TDMA cellular systems," *IEEE Trans. Commun.*, vol. 46, pp. 728-731, June 1998.

[103] M. Andersin, N.B. Mandayam and R.D. Yates, "Subspace based estimation of the signal to interterence ratio for TDMA cellular systems," *Proc. IEEE International Conf. on Vehicul. Technol. (VTC'96)*, pp. 1155-1159, April 1996.

[104] D. Ramakrishna, N.B. Mandayam, and R.D. Yates, "Subspace-based SIR estimation for CDMA cellular systems," *IEEE Trans. Vehicul. Technol.*, vol. 49, pp. 1732- 1742, Sept. 2000.

[105] H. Xu, Q. Huang and H. Zheng, "The NDA SNR estimation for QAM signals," *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Commun. (PIMRC'03)*, pp. 1162-1165, Sept. 2003.

[106] P. Gao and C. Tepedelenlioglu, "SNR estimation for non-constant modulus constellations," *Proc. IEEE Wireless Commun. and Networking Conf. (WCNC'04)*, pp. 24-29, Mar. 2004.

[107] J. Zhang, E.K.P. Chong, and D.N.C. Tse, "Output MAI distributions of linear MMSE multiuser receivers in DS-CDMA systems," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1128-1144, Mar. 2001.

[108] L.L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis.* New York: Addison-Wesley, 1991.

[109] S.S. Wilks, *Mathematical Statistics*, New York: John Wiley & Sons, 1962.

[110] B. Noble, *Applied Linear Algebra*, Engle Cliff, NJ: Prentice-Hall, 1969.

Appendix A

Proof of Independence of X and Z

A sufficient and necessary condition for the indenpendence of X and Z is that the joint characteristic function of X and Z satisfies [109, eqn. (5.3.1)]

$$\Psi_{XZ}(jv_1, jv_2) = \Psi_X(jv_1)\Psi_Z(jv_2) \tag{A.1}$$

where $\Psi_X(jv_1)$ and $\Psi_Z(jv_2)$ are the marginal characteristic functions of X and Z, respectively. Denoting $n_i = u_i + jv_i$ and $s_i^{(k)} = p_i^{(k)} + jq_i^{(k)}$ for $i = 1, 2, \cdots, I$, (2.19) and (2.21) can be rewritten as

$$X = \sum_{i=1}^{2I} x_i m_i \tag{A.2}$$

and

$$Z = \sum_{i=1}^{2I} x_i^2 - \left(\sum_{i=1}^{2I} x_i m_i\right)^2 \tag{A.3}$$

where $x_i = u_i$ and $m_i = \frac{1}{\sqrt{E_{sd}^{(k)}}} p_i^{(k)}$ when $i = 1, 2, \cdots, I$ and $x_i = v_{i-I}$ and $m_i = \frac{1}{\sqrt{E_{sd}^{(k)}}} q_{i-I}^{(k)}$ when $i = I + 1, I + 2, \cdots, 2I$. The random variables x_i, $i = 1, 2, \cdots, 2I$, are Gaussian variables each with identical PDF

$$p(x_i) = \frac{1}{\sqrt{2\pi\xi^2}} e^{-\frac{x_i^2}{2\xi^2}} \tag{A.4}$$

where $\xi^2 = \frac{N}{2}$. Denote $\mathbf{x} = (x_1 x_2 \cdots x_{2I})$. Since X is a function of $\mathbf{x}$, the marginal characteristic function of X can be derived by solving the integral $\Psi_X(jv_1) = \int e^{jv_1 X} f(\mathbf{x}) d\mathbf{x}$ where $f(\mathbf{x})$ is the joint PDF of x_i, $i = 1, 2, \cdots, 2I$, and $d\mathbf{x} = dx_1 dx_2 \cdots dx_{2I}$. This gives

$$\Psi_X(jv_1) = e^{-\frac{\xi^2}{2} v_1^2}. \tag{A.5}$$

The marginal characteristic function of Z is derived by solving the integral $\Psi_Z(jv_2) = \int e^{jv_2 Z} f(\mathbf{x}) d\mathbf{x}$. This integral can also be rewritten as

$$\Psi_Z(jv_2) = \frac{1}{(\sqrt{2\pi\xi^2})^{2l}} \int e^{-\frac{1}{2}Q_0(x_1,x_2,\cdots,x_{2l})} d\mathbf{x} \tag{A.6}$$

where $Q_0(x_1,x_2,\cdots,x_{2l}) = \sum_{i=1}^{2l}\sum_{n=1}^{2l} \gamma_{in} x_i x_n$ is a quadratic form of $\mathbf{x}$ and $\gamma_{in} = 2jv_2 m_i m_n$ for $i \neq n$ and $\gamma_{in} = \frac{1}{\xi^2} - 2jv_2 + 2jv_2 m_i^2$ for $i = n$. From [109, eqn. (7.4.3)], one has $\int e^{-\frac{1}{2}Q_0(x_1,x_2,\cdots,x_{2l})} d\mathbf{x} = \frac{(\sqrt{2\pi})^{2l}}{\sqrt{|\Gamma|}}$, where $|\Gamma|$ is the determinant of the matrix Γ whose (i,n)-th element is γ_{in}. Denote $a = 2jv_2$ and $b = \frac{1}{\xi^2} - 2jv_2$. It is derived that $|\Gamma| = b^{2l-1}(a\sum_{i=1}^{2l} m_i^2 + b) = \frac{1}{\xi^2}(\frac{1}{\xi^2} - 2jv_2)^{2l-1}$. Then, the marginal characteristic function of Z is

$$\Psi_Z(jv_2) = \frac{1}{(1 - 2jv_2\xi^2)^{\frac{2l-1}{2}}}. \tag{A.7}$$

The joint characteristic function of X and Z can also be derived by using a similar definition of $\Psi_{XZ}(jv_1, jv_2) = \int e^{jv_1 X + jv_2 Z} f(\mathbf{x}) d\mathbf{x}$. Using (A.2), (A.3) and (A.4), one has

$$\Psi_{XZ}(jv_1, jv_2) = \frac{1}{(\sqrt{2\pi\xi^2})^{2l}} \int e^{-\frac{1}{2}\left[Q_0(x_1,x_2,\cdots,x_{2l}) - 2jv_1 \sum_{i=1}^{2l} x_i m_i\right]} d\mathbf{x} \tag{A.8}$$

where $Q_0(x_1,x_2,\cdots,x_{2l})$ is the quadratic form defined as before. By completing squares for $\mathbf{x}$ in the exponent, one has $Q_0(x_1,x_2,\cdots,x_{2l}) - 2jv_1\sum_{i=1}^{2l} x_i m_i = Q_1(x_1,x_2,\cdots,x_{2l}) + \sum_{i=1}^{2l}\sum_{n=1}^{2l} \omega_{in} m_i m_n v_1^2$ where $Q_1(x_1,x_2,\cdots,x_{2l})$ is a quadratic form of $\mathbf{x}$ with complete squares and ω_{in} is the (i,n)-th element of $\Omega = \Gamma^{-1}$ [109, p. 167]. Moreover, the quadratic form $Q_1(x_1,x_2,\cdots,x_{2l})$ also satisfies $\int e^{-\frac{1}{2}Q_1(x_1,x_2,\cdots,x_{2l})} d\mathbf{x} = \frac{(\sqrt{2\pi})^{2l}}{\sqrt{|\Gamma|}}$ [109, p. 168]. Therefore, the joint characteristic function of X and Z is

$$\Psi_{XZ}(jv_1, jv_2) = \frac{1}{(1 - 2jv_2\xi^2)^{\frac{2l-1}{2}}} e^{-\frac{1}{2}\sum_{i=1}^{2l}\sum_{n=1}^{2l} \omega_{in} m_i m_n v_1^2}. \tag{A.9}$$

It is derived that $\omega_{in} = -\frac{1}{|\Gamma|} b^{2l-2} a m_i m_n$ for $i \neq n$ and $\omega_{in} = \frac{1}{|\Gamma|} b^{2l-2}(a\sum_{i=1}^{2l} m_i^2 + b - a m_i^2)$ for $i = n$. Using these in (A.9), one finally has

$$\Psi_{XZ}(jv_1, jv_2) = \frac{1}{(1 - 2j\xi^2 v_2)^{\frac{2l-1}{2}}} e^{-\frac{\xi^2}{2} v_1^2}. \tag{A.10}$$

Therefore, X and Z are independent.

Appendix B

Proof of Independence of S_1, S_2 and U

In this appendix, we prove the independence of S_1, S_2 and U defined in (2.40), (2.41) and (2.42), respectively. Again, we use the sufficient and necessary condition that [109, p. 121]

$$\Psi_{S_1 S_2 U}(jv_1, jv_2, jv_3) = \Psi_{S_1}(jv_1)\Psi_{S_2}(jv_2)\Psi_U(jv_3) \tag{B.1}$$

where $\Psi_{S_1 S_2 U}(jv_1, jv_2, jv_3)$ is the joint characteristic function of S_1, S_2 and U, $\Psi_{S_1}(jv_1)$, $\Psi_{S_2}(jv_2)$ and $\Psi_U(jv_3)$ are the marginal characteristic functions of S_1, S_2 and U, respectively. Rewrite (2.40), (2.41) and (2.42) as

$$S_1 = \sum_{i=1}^{2I} x_i m_i \tag{B.2}$$

$$S_2 = \sum_{i=1}^{2I} x_i g_i \tag{B.3}$$

$$U = \sum_{i=1}^{2I} x_i^2 - \left(\sum_{i=1}^{2I} x_i m_i\right)^2 - \left(\sum_{i=1}^{2I} x_i g_i\right)^2 \tag{B.4}$$

where x_i and m_i are the values defined as before and $g_i = -\frac{1}{\sqrt{E_{sd}^{(k)}}} q_i^{(k)}$ when $i = 1, 2, \cdots, I$ and $g_i = \frac{1}{\sqrt{E_{sd}^{(k)}}} p_{(i-I)}^{(k)}$ when $i = I+1, I+2, \cdots, 2I$. The marginal characteristic functions of S_1 and S_2 can be derived by following similar procedures to those used to derive (A.5) as

$$\Psi_{S_1}(jv_1) = e^{-\frac{\xi^2}{2} v_1^2} \tag{B.5}$$

and

$$\Psi_{S_2}(jv_2) = e^{-\frac{\xi^2}{2} v_2^2}. \tag{B.6}$$

The marginal characteristic function of U can also be shown to have similar form to that of (A.6), except that the (i,n)-th element of Γ is $\gamma_{in} = 2jv_3(m_i m_n + g_i g_n)$ for $i \neq n$ and $\gamma_{in} = \frac{1}{\xi^2} - 2jv_3 + 2jv_3 m_i^2 + 2jv_3 g_i^2$ for $i = n$ now. It is derived that $|\Gamma| = \frac{b^{2l-2}}{\xi^4}$ in this case. Therefore, $\Psi_U(jv_3)$ is found to be

$$\Psi_U(jv_3) = \frac{1}{(1-2jv_3\xi^2)^{l-1}}. \tag{B.7}$$

The joint characteristic function of S_1, S_2 and U can be derived in a similar way to that in Appendix A. From the definition of the joint characteristic function, one has

$$\Psi_{S_1S_2U}(jv_1, jv_2, jv_3) = \frac{1}{(\sqrt{2\pi\xi^2})^{2l}} \int e^{-\frac{1}{2}\left[Q_0(x_1,x_2,\cdots,x_{2l}) - 2j\sum_{i=1}^{2l} x_i(v_1 m_i + v_2 g_i)\right]} d\mathbf{x}. \tag{B.8}$$

Completing the squares for $\mathbf{x}$, one has $Q_0(x_1,x_2,\cdots,x_{2l}) - 2j\sum\limits_{i=1}^{2l} x_i(v_1 m_i + v_2 g_i) = Q_1(x_1,x_2,\cdots,x_{2l}) + \sum\limits_{i=1}^{2l}\sum\limits_{n=1}^{2l} \omega_{in}(v_1 m_i + v_2 g_i)(v_1 m_n + v_2 g_n)$ where $Q_1(x_1,x_2,\cdots,x_{2l})$ has complete squares of $\mathbf{x}$ and $\omega_{in} = \frac{b^{2l-3}}{2|\Gamma|}\left[a^2 \sum\limits_{k,h=1,k,h\neq i}^{2l} (m_k g_h - m_h g_k)^2 + 2ab \sum\limits_{k=1,k\neq i}^{2l} (m_k^2 + g_k^2)\right]$ for $i = n$, and in other cases when $i \neq n$, $\omega_{in} = -\frac{b^{2l-3}}{|\Gamma|}\left[ab(m_i m_n + g_i g_n) + a^2 \sum\limits_{k=1}^{2l} (m_i g_k - m_k g_i)(m_n g_k - m_k g_n)\right]$ now. Therefore, (B.8) is further simplified as

$$\Psi_{S_1S_2U}(jv_1, jv_2, jv_3) = \frac{1}{(1-2jv_3\xi^2)^{l-1}} e^{-\frac{1}{2}\sum_{i=1}^{2l}\sum_{n=1}^{2l} \omega_{in}(v_1 m_i + v_2 g_i)(v_1 m_n + v_2 g_n)}. \tag{B.9}$$

The summation in the exponent can be derived as $\sum\limits_{i=1}^{2l}\sum\limits_{n=1}^{2l} \omega_{in}(v_1 m_i + v_2 g_i)(v_1 m_n + v_2 g_n) = \xi^2(v_1^2 + v_2^2)$. Finally, the joint characteristic function of S_1, S_2 and U is

$$\Psi_{S_1S_2U}(jv_1, jv_2, jv_3) = \frac{1}{(1-2jv_3\xi^2)^{l-1}} e^{-\frac{\xi^2}{2}(v_1^2+v_2^2)}. \tag{B.10}$$

So, the random variables S_1, S_2 and U are independent.

Appendix C

Derivation of (2.136) and (2.150)

In this appendix, we derive (2.136) and (2.150). By using (2.132) and (2.134) in (2.133), one has

$$f(y_k, \mathbf{p}|b_k) = \frac{e^{-\frac{|y_k|^2}{2\sigma^2} - \frac{\sum_{i=\lfloor -\frac{J-1}{2} \rfloor Q}^{\lfloor \frac{J-1}{2} \rfloor Q} |p_i|^2}{2\sigma^2}}}{(4\pi^2\sigma^2)^{J+1}|\mathbf{H}_k|} \int\int e^{-\frac{1}{2}\mathbf{A}_R\mathbf{F}_k\mathbf{A}_R^T - \frac{1}{2}\mathbf{A}_I\mathbf{F}_k\mathbf{A}_I^T + \mathbf{u}\mathbf{A}_R^T + \mathbf{v}\mathbf{A}_I^T} dA_k d\mathbf{B} \qquad \text{(C.1)}$$

where $\mathbf{F}_k$, $\mathbf{A}_R$, $\mathbf{A}_I$, $\mathbf{u}$ and $\mathbf{v}$ are defined in Section 2.4. Using the fact that $\mathbf{H}_k$, $\mathbf{F}_k$ and their inverse matrices are symmetric, it can be shown that

$$-\frac{1}{2}\mathbf{A}_R\mathbf{F}_k\mathbf{A}_R^T + \mathbf{u}\mathbf{A}_R^T = -\frac{1}{2}\mathbf{A'}_R\mathbf{F}_k\mathbf{A'}_R^T + \frac{1}{2}\mathbf{u}\mathbf{F}_k^{-1}\mathbf{u}^T \qquad \text{(C.2)}$$

and

$$-\frac{1}{2}\mathbf{A}_I\mathbf{F}_k\mathbf{A}_I^T + \mathbf{v}\mathbf{A}_I^T = -\frac{1}{2}\mathbf{A'}_I\mathbf{F}_k\mathbf{A'}_I^T + \frac{1}{2}\mathbf{v}\mathbf{F}_k^{-1}\mathbf{v}^T \qquad \text{(C.3)}$$

where $\mathbf{A'}_R = \mathbf{A}_R - \mathbf{u}\mathbf{F}_k^{-1}$ and $\mathbf{A'}_I = \mathbf{A}_I - \mathbf{v}\mathbf{F}_k^{-1}$. Putting (C.2) and (C.3) in (C.1) and executing a transformation of variables, one can obtain

$$f(y_k, \mathbf{p}|b_k) = \frac{e^{-\frac{|y_k|^2}{2\sigma^2} - \frac{\sum_{i=\lfloor -\frac{J-1}{2} \rfloor Q}^{\lfloor \frac{J-1}{2} \rfloor Q} |p_i|^2}{2\sigma^2} + \frac{1}{2}\mathbf{u}\mathbf{F}_k^{-1}\mathbf{u}^T + \frac{1}{2}\mathbf{v}\mathbf{F}_k^{-1}\mathbf{v}^T}}{(4\pi^2\sigma^2)^{J+1}|\mathbf{H}_k|} \int\int e^{-\frac{1}{2}\mathbf{A'}_R\mathbf{F}_k\mathbf{A'}_R^T - \frac{1}{2}\mathbf{A'}_I\mathbf{F}_k\mathbf{A'}_I^T} d\mathbf{A'}_R d\mathbf{A'}_I. \qquad \text{(C.4)}$$

Note that the random variables, $\mathbf{A'}_R$ and $\mathbf{A'}_I$, are Gaussian since $\mathbf{A}_R$ and $\mathbf{A}_I$ are jointly Gaussian and the transformations are linear. Therefore, they satisfy [109, eqn. (7.4.3)]

$$\int e^{-\frac{1}{2}\mathbf{A'}_R\mathbf{F}_k\mathbf{A'}_R^T} d\mathbf{A'}_R = \int e^{-\frac{1}{2}\mathbf{A'}_I\mathbf{F}_k\mathbf{A'}_I^T} d\mathbf{A'}_I = \frac{(2\pi)^{\frac{J+1}{2}}}{|\mathbf{F}_k|^{\frac{1}{2}}}. \qquad \text{(C.5)}$$

Substituting (C.5) in (C.4), and after doing some mathematical manipulations, (2.136) can be obtained. Equation (2.150) can be derived in a similar way.

Appendix D

Derivation of $\mathbf{S}_k$ and $\mathbf{Q}_k$

Here, we derive the expressions for $\mathbf{S}_k$ in (2.139) and $\mathbf{Q}_k$ in (2.153). Since $\mathbf{F}_k = \mathbf{H}_k^{-1} + \mathbf{G}_k$, by using [110, eqn. (5.32)], one has

$$\mathbf{F}_k^{-1} = \mathbf{H}_k - \mathbf{H}_k \cdot (\mathbf{G}_k^{-1} + \mathbf{H}_k)^{-1} \mathbf{H}_k. \tag{D.1}$$

It can be shown that

$$\mathbf{G}_k^{-1} + \mathbf{H}_k = \begin{bmatrix} \frac{\sigma^2}{|b_k|^2} + E_p^2\alpha^2 & \mathbf{w}_k \\ \mathbf{w}_k^T & \frac{\sigma^2}{|\bar{b}|^2}\mathbf{E} + \mathbf{C}_k \end{bmatrix}. \tag{D.2}$$

Therefore, the inverse of $\mathbf{G}_k^{-1} + \mathbf{H}_k$ is obtained from (D.2) as [110, eqn. (1.35)]

$$(\mathbf{G}_k^{-1} + \mathbf{H}_k)^{-1} = \begin{bmatrix} \frac{|b_k|^2}{\sigma^2 + E_p^2\alpha^2|b_k|^2}\left(1 + \frac{|b_k|^2 \mathbf{w}_k \mathbf{Z}_1 \mathbf{w}_k^T}{\sigma^2 + E_p^2\alpha^2|b_k|^2}\right) & -\frac{|b_k|^2}{\sigma^2 + E_p^2\alpha^2|b_k|^2}\mathbf{w}_k \mathbf{Z}_1 \\ -\frac{|b_k|^2}{\sigma^2 + E_p^2\alpha^2|b_k|^2}\mathbf{Z}_1 \mathbf{w}_k^T & \mathbf{Z}_1 \end{bmatrix} \tag{D.3}$$

where $\mathbf{Z}_1$ is defined as before. Substituting (2.135) and (D.3) into (D.1) and performing the matrix multiplication, one has

$$\mathbf{F}_k^{-1} = \begin{bmatrix} \frac{E_p^2\alpha^2\sigma^2}{\sigma^2 + E_p^2\alpha^2|b_k|^2} - \frac{\sigma^4 \mathbf{w}_k \mathbf{Z}_1 \mathbf{w}_k^T}{(\sigma^2 + E_p^2\alpha^2|b_k|^2)^2} & \frac{\sigma^2(\mathbf{w}_k - \mathbf{w}_k \mathbf{Z}_1 \mathbf{C}_k)}{\sigma^2 + E_p^2\alpha^2|b_k|^2} + \frac{\sigma^2|b_k|^2 \mathbf{w}_k \mathbf{Z}_1 \mathbf{w}_k^T \mathbf{w}_k}{(\sigma^2 + E_p^2\alpha^2|b_k|^2)^2} \\ \frac{\sigma^2(\mathbf{w}_k^T - \mathbf{C}_k \mathbf{Z}_1 \mathbf{w}_k^T)}{\sigma^2 + E_p^2\alpha^2|b_k|^2} + \frac{\sigma^2|b_k|^2 \mathbf{w}_k \mathbf{Z}_1 \mathbf{w}_k^T \mathbf{w}_k^T}{(\sigma^2 + E_p^2\alpha^2|b_k|^2)^2} & \mathbf{Z}_3 \end{bmatrix} \tag{D.4}$$

where $\mathbf{Z}_3 = \mathbf{C}_k - \frac{|b_k|^2 \mathbf{w}_k^T \mathbf{w}_k}{\sigma^2 + E_p^2\alpha^2|b_k|^2}\left(1 + \frac{|b_k|^2 \mathbf{w}_k \mathbf{Z}_1 \mathbf{w}_k^T}{\sigma^2 + E_p^2\alpha^2|b_k|^2}\right) + \frac{|b_k|^2(\mathbf{C}_k \mathbf{Z}_1 \mathbf{w}_k^T \mathbf{w}_k + \mathbf{w}_k^T \mathbf{w}_k \mathbf{Z}_1 \mathbf{C}_k)}{\sigma^2 + E_p^2\alpha^2|b_k|^2} - \mathbf{C}_k \mathbf{Z}_1 \mathbf{C}_k$. Since $\mathbf{S}_k$ is the first row of $\mathbf{F}_k^{-1}$ excluding the first element, (2.139) can be obtained from (D.4). Also, one has $\mathbf{F}_k \mathbf{H}_k = \begin{bmatrix} \frac{E_p^2\alpha^2|b_k|^2}{\sigma^2} + 1 & \frac{|b_k|^2}{\sigma^2}\mathbf{w}_k \\ \frac{|\bar{b}|^2}{\sigma^2}\mathbf{w}_k^T & \frac{|\bar{b}|^2}{\sigma^2}\mathbf{C}_k + \mathbf{E} \end{bmatrix}$. The inverse of $\mathbf{F}_k \mathbf{H}_k$ can be obtained by using [110, eqn.

(1.35)]. This gives

$$(\mathbf{F}_k\mathbf{H}_k)^{-1} = \begin{bmatrix} \frac{\sigma^2}{\sigma^2+E_p^2\alpha^2|b_k|^2}\left(1+\frac{|b_k|^2|\bar{b}|^2\mathbf{w}_k\mathbf{Z}_4\mathbf{w}_k^T}{\sigma^2(\sigma^2+E_p^2\alpha^2|b_k|^2)}\right) & -\frac{|b_k|^2}{\sigma^2+E_p^2\alpha^2|b_k|^2}\mathbf{w}_k\mathbf{Z}_4 \\ -\frac{|\bar{b}|^2}{\sigma^2+E_p^2\alpha^2|b_k|^2}\mathbf{Z}_4\mathbf{w}_k^T & \mathbf{Z}_4 \end{bmatrix} \tag{D.5}$$

where $\mathbf{Z}_4 = \frac{\sigma^2}{|\bar{b}|^2}\mathbf{Z}_1$ and $\mathbf{Z}_1$ is defined as before. Finally, since $\mathbf{Q}_k$ is the transpose of the first column of $(\mathbf{F}_k\mathbf{H}_k)^{-1}$, one can obtain (2.153) from (D.5).

Appendix E

Proof of the Equivalence between (2.144) and (2.163)

We prove the equivalence between (2.144) and (2.163) here. Denote

$$\mathbf{R}_k = \left(1 + \frac{|b_k|^2 \mathbf{w}_k \mathbf{Z}_1 \mathbf{w}_k^T}{\sigma^2 + E_p^2 \alpha^2 |b_k|^2}\right) \mathbf{E} - \mathbf{Z}_1 \cdot \mathbf{C}_k. \tag{E.1}$$

From (2.139), (2.161), (2.144) and (2.163), it is enough to show that

$$\mathbf{w}_k \mathbf{R}_k^T \mathbf{w}_k^T = \sqrt{\mathbf{w}_k (|\tilde{b}|^2 \mathbf{C}_k + \sigma^2 \mathbf{E})^{-1} \mathbf{w}_k^T \cdot \mathbf{w}_k \mathbf{R}_k (|\tilde{b}|^2 \mathbf{C}_k + \sigma^2 \mathbf{E}) \mathbf{R}_k^T \mathbf{w}_k^T}. \tag{E.2}$$

Using (2.140) and [110, eqn. (5.32)], one has

$$\mathbf{w}_k^T \mathbf{w}_k \mathbf{Z}_1 (|\tilde{b}|^2 \mathbf{C}_k + \sigma^2 \mathbf{E}) = |\tilde{b}|^2 \left(1 + \frac{|b_k|^2 \mathbf{w}_k \mathbf{Z}_1 \mathbf{w}_k^T}{\sigma^2 + E_p^2 \alpha^2 |b_k|^2}\right) \mathbf{w}_k^T \mathbf{w}_k. \tag{E.3}$$

This in turn gives

$$\begin{aligned} \mathbf{w}_k^T \mathbf{w}_k \mathbf{Z}_1 (|\tilde{b}|^2 \mathbf{C}_k + \sigma^2 \mathbf{E}) \mathbf{C}_k - |\tilde{b}|^2 \left(1 + \frac{|b_k|^2 \mathbf{w}_k \mathbf{Z}_1 \mathbf{w}_k^T}{\sigma^2 + E_p^2 \alpha^2 |b_k|^2}\right) \mathbf{w}_k^T \mathbf{w}_k \mathbf{C}_k = \mathbf{0} \\ \mathbf{C}_k (|\tilde{b}|^2 \mathbf{C}_k + \sigma^2 \mathbf{E}) \mathbf{Z}_1 \mathbf{w}_k^T \mathbf{w}_k - |\tilde{b}|^2 \left(1 + \frac{|b_k|^2 \mathbf{w}_k \mathbf{Z}_1 \mathbf{w}_k^T}{\sigma^2 + E_p^2 \alpha^2 |b_k|^2}\right) \mathbf{C}_k \mathbf{w}_k^T \mathbf{w}_k = \mathbf{0} \end{aligned} \tag{E.4}$$

where $\mathbf{0}$ is a $J \times J$ zero matrix. From (E.4), it is easy to verify that

$$\begin{aligned} & \mathbf{w}_k^T \mathbf{w}_k \left[\left(1 + \frac{|b_k|^2 \mathbf{w}_k \mathbf{Z}_1 \mathbf{w}_k^T}{\sigma^2 + E_p^2 \alpha^2 |b_k|^2}\right) (|\tilde{b}|^2 \mathbf{C}_k + \sigma^2 \mathbf{E}) - \mathbf{Z}_1 (|\tilde{b}|^2 \mathbf{C}_k + \sigma^2 \mathbf{E}) \mathbf{C}_k\right] \\ &= \left[\left(1 + \frac{|b_k|^2 \mathbf{w}_k \mathbf{Z}_1 \mathbf{w}_k^T}{\sigma^2 + E_p^2 \alpha^2 |b_k|^2}\right) (|\tilde{b}|^2 \mathbf{C}_k + \sigma^2 \mathbf{E}) - \mathbf{C}_k (|\tilde{b}|^2 \mathbf{C}_k + \sigma^2 \mathbf{E}) \mathbf{Z}_1\right] \mathbf{w}_k^T \mathbf{w}_k. \end{aligned} \tag{E.5}$$

Since $(|\tilde{b}|^2 \mathbf{C}_k + \sigma^2 \mathbf{E}) \mathbf{C}_k = \mathbf{C}_k (|\tilde{b}|^2 \mathbf{C}_k + \sigma^2 \mathbf{E})$, from (E.5), one further has

$$\mathbf{w}_k^T \mathbf{w}_k \mathbf{R}_k (|\tilde{b}|^2 \mathbf{C}_k + \sigma^2 \mathbf{E}) = (|\tilde{b}|^2 \mathbf{C}_k + \sigma^2 \mathbf{E}) \mathbf{R}_k^T \mathbf{w}_k^T \mathbf{w}_k \tag{E.6}$$

where $\mathbf{R}_k$ is given by (E.1). Multiplying both sides of (E.6) with $\mathbf{w}_k(|\tilde{b}|^2\mathbf{C}_k + \sigma_n^2\mathbf{E})^{-1}$ on the left and $\mathbf{R}_k^T\mathbf{w}_k^T$ on the right, one has

$$\mathbf{w}_k(|\tilde{b}|^2\mathbf{C}_k + \sigma^2\mathbf{E})^{-1}\mathbf{w}_k^T \cdot \mathbf{w}_k\mathbf{R}_k(|\tilde{b}|^2\mathbf{C}_k + \sigma^2\mathbf{E})\mathbf{R}_k^T\mathbf{w}_k^T = \mathbf{w}_k\mathbf{R}_k^T\mathbf{w}_k^T\mathbf{w}_k\mathbf{R}_k^T\mathbf{w}_k^T. \tag{E.7}$$

From (E.7), eqn. (E.2) can be obtained. Therefore, (2.144) and (2.163) are equivalent.